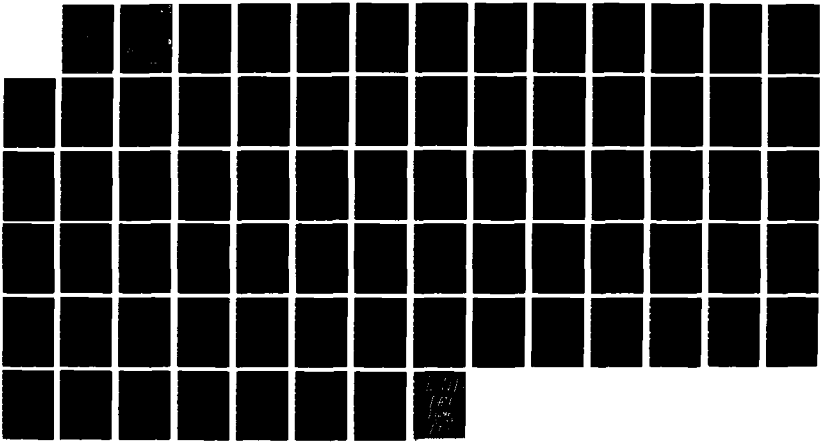
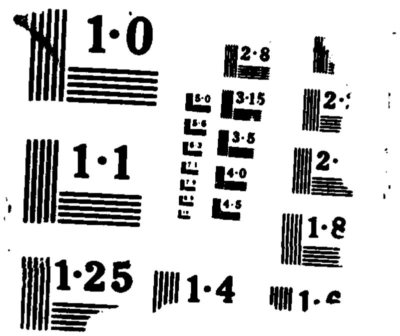
AD-A193 318   ROBUST PREDICTION FOR STATIONARY PROCESSES 2D ENRICHED   1/1
VESSION (U) VIRGINIA UNIV CHARLOTTESVILLE DEPT OF
ELECTRICAL ENGINEERING   P PAPANTONI-KAZAKOS ET AL.
UNCLASSIFIED 24 NOV 87 UVA/525682/EE88/103   F/G 12/3   NL




1·0
2·8
3·15
3·5
4·0
4·5
1·1
1·8
1·25
1·4

AD-A193 318

AFOSR-TR- 88-0361

(2)

[redacted]

Grant No. AFOSR-87-0224
July 1, 1987 - June 30, 1988

DTIC FILE COPY

ROBUST PREDICTION FOR STATIONARY PROCESSES
2d Enriched Version

Submitted to:

Air Force Office of Scientific Research
Building 410
Bolling Air Force Base
Washington, D.C. 20332-6448

Attention: Major Brian W. Woodruff, NM

Submitted by:

P. Papantoni-Kazakos
Department of Electrical Engineering
University of Virginia
Charlottesville, Virginia 22901

Kailash Birmiwal
Department of Electrical Engineering
Southern Illinois University
Carbondale, Illinois 62901

Report No. UVA/525682/EE88/103
November 1987

DTIC
ELECTE
APR 01 1988
S D
E

SCHOOL OF ENGINEERING AND APPLIED SCIENCE

DEPARTMENT OF ELECTRICAL ENGINEERING

This document has been approved
for public release and sale; its
distribution is unlimited.

UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA 22901

88 3 31 020

[redacted]

Grant No. AFOSR-87-0224
July 1, 1987 - June 30, 1988

ROBUST PREDICTION FOR STATIONARY PROCESSES
2d Enriched Version

Submitted to:

Air Force Office of Scientific Research
Building 410
Bolling Air Force Base
Washington, D.C. 20332-6448

Attention: Major Brian W. Woodruff, NM

Submitted by:

P. Papantoni-Kazakos
Department of Electrical Engineering
University of Virginia
Charlottesville, Virginia 22901

Kailash Birmiwal
Department of Electrical Engineering
Southern Illinois University
Carbondale, Illinois 62901

Department of Electrical Engineering
SCHOOL OF ENGINEERING AND APPLIED SCIENCE
UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA

Accession For	
NTIS GRA&I	X
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

INSPECTED 4

Report No. UVA/525682/EE88/103

Copy No. 4

November 1987

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: Unclassified	1b. RESTRICTIVE MARKINGS: None
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for Public Release; Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): UVA/525682/EE88/103	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 88-0361
6a. NAME OF PERFORMING ORGANIZATION: University of Virginia, Dept. of Electrical Engr. / 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research/PKZ
6c. ADDRESS (City, State, and ZIP Code): Thornton Hall, Charlottesville, VA 22901	7b. ADDRESS (City, State, and ZIP Code): Building 410, Bolling Air Force Base, Washington, D.C. 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Air Force Office of Scientific Research/NM / 8b. OFFICE SYMBOL (If applicable): NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-87-0224
8c. ADDRESS (City, State, and ZIP Code): Building 410, Bolling Air Force Base, Washington, D.C. 20332-6448	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. 61102F; PROJECT NO. 2304; TASK NO. A6; WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
Robust Prediction For Stationary Processes 2d Enriched Version

12. PERSONAL AUTHOR(S)
P. Papantoni-Kazakos

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
~~Annual Technical~~ Journal	FROM 07/01/87 TO 06/30/88	87 November 24	69

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

We consider prediction for stationary processes, in environments where data outliers may be present. ~~We develop~~ two sequences ~~of~~ outlier resistant prediction operations, which are uniformly qualitatively robust. ~~We study~~ the asymptotic mean-squared performance of the developed operations in the absence of data outliers. Important performance characteristics studied include the breakdown point and the influence function. We include numericl results, for some autoregressive nominal processes.

we developed for

are studied

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☐ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: Unclassified
22a. NAME OF RESPONSIBLE INDIVIDUAL: Major Brian W. Woodruff	22b. TELEPHONE (Include Area Code): (202) 767-5025 / 22c. OFFICE SYMBOL: NM

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

TABLE OF CONTENTS

II. PRELIMINARIES

Let R be the real line, and let B be the usual Borel σ-field on R. Let R^∞ be the two-sided sequence space, and let B^∞ be the Borel σ-field on R^∞ that is generated by the product topology on R^∞. We consider a real-valued discrete-time process, $\{X_n, \infty<n<\infty\}$, whose measure μ_o is known and is defined on B^∞. We name $\{x_n, -\infty<n<\infty\}$ the <u>nominal process</u> , and we denote by $\{x_n, -\infty<n<\infty\}$ data realizations generated by it. Let $\hat{x}_n = \hat{x}_n(x_1^{n-1})$ denote the optimal one-step mean-squared prediction operation, given the sequence realization $x_1^{n-1} \overset{\Delta}{=} \{x_l, 1\le l\le n-1\}$, when $\{x_n, -\infty<n<\infty\}$ is generated by the nominal process. Then, if $g_n = g_n(x_1^{n-1})$ denotes some scalar real-valued function on the sequence x_1^{n-1}, we have:

$$e_n(\mu_o, \hat{x}_n) = \inf_{g_n} e_n(\mu_o, g_n) \tag{1}$$

$$\hat{x}_n(x_1^{n-1}) = E_{\mu_o}\{X_n \mid x_1^{n-1}\} \tag{2}$$

; where $E_{\mu_o}\{\ \}$ denotes expectation with respect to the measure μ_o, where $X_1^n \overset{\Delta}{=} \{X_l, 1\le l\le n\}$, and where,

$$e_n(\mu_o, g_n) \overset{\Delta}{=} E_{\mu_o}\{[X_n - g_n(X_1^{n-1})]^2\} \tag{3}$$

The expression in (3) is called the one-step prediction error induced by g_n at μ_o. Let L_n denote the class of all the scalar real-valued linear functions defined on R^n. Let then $\hat{x}_n^L = \hat{x}_n^L(x_1^{n-1})$ be such that:

$$e_n(\mu_o, \hat{x}_n^L) = \inf_{g_n^L \varepsilon L_{n-1}} e_n(\mu_o, g_n^L) \tag{4}$$

Then, $\hat{x}_n^L$ is called the optimal linear one-step mean squared predictor at μ_o, given the sequence realization x_1^{n-1}, and generally,

$$e_n(\mu_o, \hat{x}_n) \le e_n(\mu_o, \hat{x}_n^L) \tag{5}$$

If the measure μ_o is Gaussian, then $\hat{x}_n(x_1^{n-1}) \overset{\Delta}{=} \hat{x}_n^L(x_1^{n-1}), \forall n$, and (5) is then satisfied with equality for all n. If μ_o is non Gaussian, then (5) is generally a strict inequality.

The above summary corresponds to parametric one-step prediction; that is, it corresponds to the case where the measure μ_o that generates the data sequences is known. In this paper, we are concerned with the outlier model. Then, the observation process $\{Y_n, -\infty<n<\infty\}$ is generated by three mutually independent processes; the nominal process $\{X_n, -\infty<n<\infty\}$ and two i.i.d. processes $\{V_n, -\infty<n<\infty\}$ and $\{Z_n, -\infty<n<\infty\}$, as follows:

$$Y_n = (1-V_n)X_n + V_n Z_n, \quad n=\ldots-1,0,1,\ldots \tag{6}$$

; where the common distribution of the variables Z_n is unknown, and where $\{V_n, -\infty<n<\infty\}$ is a binary process. In particular, for some given ε : $0\le\varepsilon<1$, the latter process is such that:

$$P(V_k = 0) = 1-\varepsilon$$

$$P(V_k = 1) = \varepsilon \tag{7}$$

In the outlier model in (6), $\{Z_n, -\infty<n<\infty\}$ is called the <u>contaminating process</u>, and $\{V_n, -\inf<n<\infty\}$ determines the <u>contamination law</u>. In the presence of the latter model, the objective is prediction of the nominal datum x_n, given the observation sequence y_1^{n-1}, for all n, and the problem formalization is then clearly nonparametric. Let μ denote the measure of the observation process, and let $\{g_n\}_{2\le n<\infty}$ denote a sequence of one-step predictors, where $g_n = g_n(y_1^{n-1})$. Let us then define,

$$e_n(\mu,g_n) \stackrel{\Delta}{=} E_\mu\{[X_n-g_n(Y_1^{n-1})]^2\} \tag{8}$$

In (8), $e_n(\mu, g_n)$ is the mean-squared error induced by the predictor g_n, when the measure of the observation process $\{Y_n, -\infty<n<\infty\}$ is μ, and where X_n is generated by the nominal process whose measure is μ_o. Clearly, $e_n(\mu_o, g_n)$ is then as in (3), and it represents the mean-squared performance of the predictor g_n at the nominal measure μ_o, (that is, when outliers are absent).

Our objective is to design a sequence $\{g_n\}_{2\leq n<\infty}$ of predictors whose mean-squared performance is stable in the presence of variations in the measure μ of the observation process $\{Y_n, -\infty<n<\infty\}$. This stability corresponds to qualitative robustness, and is defined as follows:

Given $\eta>0$, there exists $\delta>0$, such that:

$$\Pi_\rho(\mu_o, \mu)<\delta \text{ implies } |e_n(\mu_o, g_n)-e_n(\mu, g_n)|<\eta\ ;\ \forall n$$

In the above definition, Π_ρ denotes Prohorov distance with an appropriate distortion measure ρ on data sequences, and sequences $\{g_n\}$ of operations that satisfy this stability are called <u>qualitatively robust</u> at the measure μ_o. As found first in [13], and later in [1],[14], and [16], for the sequence $\{g_n\}$ to be qualitatively robust, pointwise continuity and asymptotic continuity in conjunction with boundness, are sufficient. In particular, it is sufficient that: g_n is bounded for all n, and:

(A) Given finite n, given $\eta>0$, given x_1^n, there exists $\delta>0$, such that, $y_1^n : \gamma_n(x_1^n, y_1^n) \stackrel{\Delta}{=} n^{-1}\sum_{i=1}^{n} |x_i-y_i|<\delta$ implies $|g_{n+1}(x_1^n)-g_{n+1}(y_1^n)|<\eta$.

(B) Given μ_o stationary, given $\zeta>0$, $\eta>0$, there exist integers n_o, m, some $\delta>0$, and for each $n>n_o$ some $\Delta^n \varepsilon R^n$ with $\mu_o(\Delta^n)>1-\eta$, such that for each $x^n \varepsilon \Delta^n$ and y^n such that $\inf\{\alpha : \#[i : \gamma_m(x_i^{i+m-1}, y_i^{i+m-1})>\alpha] \leq n\alpha\} < \delta$ it is implied that $|g_{n+1}(x_1^n)-g_{n+1}(y_1^n)| < \zeta$.

Given a sequence $\{g_n\}$ of predictors which is qualitatively robust at the nominal measure μ_o, its important quantitative performance criteria are: (1) Its asymptotic mean-squared performance at the nominal measure, $\limsup_{n\to\infty} e_n(\mu_o, g_n)$. (2) Its breakdown point. (3) Its influence function. The breakdown point and the influence function represent measures of resistance to outliers, and their definitions are given below.

Consider the model in (6), and let then $\{Z_n\}$ be a deterministic process with amplitude w; that is, $P(Z_n = w) = 1$. Let then $\mu_{\varepsilon,w}$ be the measure of the observation process $\{Y_n\}$. Given a sequence $\{g_n\}$ of predictors, we then define:

<u>Influence Function</u> of the sequence $\{g_n\}$:

$$I_g(w) \overset{\Delta}{=} \lim_{\varepsilon\to 0} \frac{e(\mu_{\varepsilon,w}, g) - e(\mu_o, g)}{\varepsilon} \tag{9}$$

; where,

$$e(\mu,g) \overset{\Delta}{=} \limsup_{n\to\infty} e_n(\mu, g_n) \tag{10}$$

provided the limit in (9) exists.

<u>Breakdown Point</u> of the sequence $\{g_n\}$:

$$\varepsilon_g^* \overset{\Delta}{=} \sup\{\varepsilon : \sup e(\mu_{\varepsilon,w}, g) \le \limsup_{n\to\infty} E_{\mu_o}\{X_n^2\}\} \tag{11}$$

; where $e(\mu, g)$ is defined as in (10).

We note that the breakdown point is the maximum frequency of independent outliers that the prediction sequence can tolerate asymptotically, without becoming useless, (that is, before the observation sequences provide no information about the next process datum), where the amplitude of the outliers is arbitrary. Alternatively, we can define the breakdown point as

$$\bar{\varepsilon}_g^* \stackrel{\Delta}{=} \sup\{\varepsilon:\ \limsup_{w\to\infty} e(\mu_{\varepsilon,w}, g) \le \limsup_{n\to\infty} E_{\mu_o}\{X_n^2\}\} \tag{12}$$

If $e(\mu_{\varepsilon,w}, g)$ is symmetric in w about zero and is monotonically increasing in $|w|$, then $\bar{\varepsilon}_g^* = \varepsilon_g^*$. In general, ε_g^* is defined in terms of a stronger condition than $\bar{\varepsilon}_g^*$ and hence

$$\varepsilon_g^* \le \bar{\varepsilon}_g^* \tag{13}$$

The influence function represents the slope of the function $e(\mu_{\varepsilon,w}, g) - e(\mu_o,g) \stackrel{\Delta}{=} F_{\varepsilon,g}(w)$, at the $\varepsilon=0$ point. $F_{\varepsilon,g}(w)$ corresponds to the asymptotic mean-squared error increase induced by the prediction sequence $\{g_n\}$, when from absence of outliers the environment shifts to ε-frequency and w-amplitude outlier occurrence.

The outler model in (6) can be generalized to i.i.d. sequences of m-size blocks of outliers, as follows:

$$Y_{(k-1)m+1}^{km} = (1-V_k)X_{(k-1)m+1}^{km} + V_k Z_{(k-1)m+1}^{km}\ ;\ \ k = \ldots,-1,0,1,\ldots \tag{14}$$

; where the sequence $\{V_k\}$ is as in (7), and where the vector random variables $\{Z_{(k-1)m+1}^{km}\}$ are i.i.d. with unknown distribution. Let $\mu_{\varepsilon,w,m}$ denote the measure of the observation process $\{Y_n\}$, when the model in (14) is present, and when $P(Z_n = w) = 1$. Then, given a sequence $\{g_n\}$ of predictors, and defining $e(\mu,g)$ as in (10), the breakdown point, $\varepsilon_{g,m}^*$, and the influence function, $I_{g,m}(w)$, that correspond to the outlier model in (14) are defined as follows:

$$\varepsilon_{g,m}^* \stackrel{\Delta}{=} \sup\{\varepsilon : \sup_w e(\mu_{\mu,w,m}, g) \le \limsup_{n\to\infty} E_{\mu_o}\{X_n^2\}\} \tag{15}$$

$$I_{g,m}(w) \overset{\Delta}{=} \lim_{\varepsilon \to 0} \frac{e(\mu_{\varepsilon,w,m}, g) - e(\mu_o, g)}{\varepsilon} \tag{16}$$

provided the limit in (16) exists. We also define $\bar{\varepsilon}^*_{g,m}$ by replacing supremum over w with $\lim$ sup as in the case of (12).

III. OUTLIER RESISTANT PREDICTION OPERATIONS

We consider a stationary, zero mean, real-valued process $\{X_n, -\infty<n<\infty\}$, with measure μ_o, and $E_{\mu_o}\{X_n^2\} = \sigma^2 < \infty$. We also consider the outlier model in (14) for the observation process $\{Y_n, -\infty<n<\infty\}$. We concentrate on the design of qualitatively robust and outlier resistant sequences $\{g_n\}$ of one-step predictors for the process $\{X_n, -\infty<n<\infty\}$. Our methodology involves two steps: (1) A saddle-point game formalization and solution for the predictors $g_n : 2\leq n\leq m+1$. (2) A qualitatively robust generalization of the solutions in Step 1; for the predictors $g_n : n>m+1$.

In the sequel, we will assume that both the nominal and the contaminating processes are absolutely continous. We will then denote by $f_o(x_1^n)$ the density function induced by the nominal process at the vector point x_1^n; we will denote by $f(y_1^n)$ the density function induced by the observation process at the vector point y_1^n. We note that then, for n : $2\leq n\leq m+1$, the class F_n, of density functions induced by the model in (14) is as follows:

$$F_n = \{f : f(y_1^{n-1}) - (1-\varepsilon)f_o(y_1^{n-1}) \geq 0 : \forall y_1^{n-1} \varepsilon R^{n-1}, \int_{R^{n-1}} f(y_1^{n-1})dy_1^{n-1} = 1\} \tag{17}$$

<u>Construction of Prediction Operations - Step 1</u>

Let us consider the model in (14) and one-step prediction based on observation sequences y_1^{n-1}, with n : $2\leq n\leq m+1$. Given such an n, we consider the following saddle point game, where F_n is as in (17):

Find a pair, (f^*, g_n^*), of an observation density function and an one-step predictor, such that $f^* \varepsilon F_n$, and:

$$\forall f \varepsilon F_n : c_n(f, g_n^*) \leq c_n(f^*, g_n^*) \leq c_n(f^*, g_n) ; \forall g_n \tag{18}$$

In (18), the errors $c_n(f, g_n)$ are as in (8), where the measure, μ, has been substituted by the corresponding density function, f.

Consider a pair, (f', g_n'), of an observation density and a prediction operation, such that:

$$(f', g_n') : c_n(f', g_n') = \sup_{f \varepsilon F_n} \inf_{g_n} c_n(f, g_n) \tag{19}$$

From the results in [15] we then conclude that if the operation $g_n' = g_n'(y_1^{n-1})$ is pointwise continuous and bounded, then $(f', g_n') \equiv (f^*, g_n^*)$, and the pair is a unique solution of the game in (18). We now present a theorem whose proof is in the Appendix.

Theorem 1 Let the nominal process be zero mean Gaussian. Let $m_o(y_1^{n-1}) = B_{n-1}^T y_1^{n-1}$ denote the optimal at the Gaussian process one-step predictor, when the observation sequence is y_1^{n-1}. Let $n : 2 \leq n \leq m+1$. Then, the pair (f', g_n') in (19) is as follows:

$$g_n'(y_1^{n-1}) = m_o(y_1^{n-1}) \cdot \min(1, \lambda_{n-1} | m_o(y_1^{n-1}) |^{-1}) \tag{20}$$

$$f'(y_1^{n-1}) = (1-\varepsilon) f_o(y_1^{n-1}) \cdot \max(1, \lambda_{n-1}^{-1} | m_o(y_1^{n-1}) |) \tag{21}$$

; where λ_{n-1} is unique, and such that:

$$\lambda_{n-1} : \int_{R^{n-1}} f'(y_1^{n-1}) dy_1^{n-1} = 1 \tag{22}$$

Since the operation in (20) is pointwise continuous and bounded, $(f', g_n') \equiv (f^*, g_n^*)$, and the pair is a unique solution of the game in (18).

When the nominal process is non Gaussian, the operation g_n' in (19) is generally not pointwise continuous; thus, there is no guarantee then that it will satisfy the game in (18), and it is generally qualitatively nonrobust. However, drawing from linear prediction in the absence of outliers, we will adopt the operations in Theorem 1, for non Gaussian nominal proceses as well. Specifically:

Let the nominal process be zero mean stationary. Let $m_o(y_1^{n-1}) = B_{n-1}^T y_1^{n-1}$ denote the optimal at the nominal process linear one-step predictor when the observation sequence is y_1^{n-1}. Let f_G denote density of the Gaussian process whose power spectral density is the same as that of the nominal process, and whose mean is zero. Then, in the presence of the model in (14), and for $n: 2 \le n \le m+1$, we adopt the following one step prediction operation:

$$g_n^*(y_1^{n-1}) = m_o(y_1^{n-1}) \cdot \min(1, \lambda_{n-1} | m_o(y_i^{n-1}) |^{-1})$$

$$\lambda_{n-1} : \int_{R^{n-1}} f_G(y_1^{n-1}) \cdot \max(1, \lambda_{n-1}^{-1} | m_o(y_1^{n-1}) |) = (1-\varepsilon)^{-1} \tag{23}$$

We note that for $\varepsilon=0$, the value of λ_{n-1} is infinity and the operation g_n^* becomes identical to the optimal of the nominal linear one-step predictor. As ε increases, λ_{n-1} decreases monotonically, becoming zero at $\varepsilon=1$.

Construction of Prediction Operations - Step 2

In this part, we are concerned with the construction of qualitatively robust prediction operations, for large dimensionalities of observation sequences. We point out that the operations in (23) are qualitatively robust for finite such dimensionalities only. Indeed, they satisfy condition (A) in section II and are bounded, but they do not satisfy condition (B). At the same time, the outlier model in (12) does not allow for the formalization of a saddle point game for arbitrary data dimensionalities, even when the nominal process is Gaussian. We will thus adopt an ad hoc approach.

Let $\{a_j^{(n)}\}_{1 \le j \le n}$ denote the one-step prediction coefficients of the nominal process, when n observation data are available. That is, if $m_o(y_1^n)$ denotes the optimal at the nominal linear one-step predictor when the observation sequence is y_1^n, then:

$$m_o(y_1^n) = \sum_{j=1}^{n} a_j^{(n)} y_j \tag{24}$$

Let g_n^* be as in (23). Then, we propose the following sequences, $\{G_{1,n}^*\}$ and $\{G_{2,n}^*\}$ of one-step predictors:

Sequence $\{G_{1,n}^*\}$

$$G_{1,n}^*(y_1^{n-1}) = G_{1,n}^{*m}(y_1^{n-1}) = g_n^*(y_1^{n-1}) \quad ; \quad \text{for } 2 \le n \le m+1$$

$$G_{1,n}^*(y_1^{n-1}) = G_{1,n}^{*m}(y_1^{n-1}) = \sum_{j=1}^{m} a_j^{(n-1)} \left\{ \frac{g_{j+1}^*(y_1^j) - g_{j+1}^*(y_1^{j-1}, 0)}{a_j^{(j)}} \right\}$$

$$+ \sum_{j=m+1}^{n-1} a_j^{(n-1)} \left\{ \frac{g_{m+1}^*(y_{j-m}^{j}) - g_{m+1}^*(y_{j-m+1}^{j-1}, 0)}{a_m^{(m)}} \right\}, \quad \text{for } n > m+1 \tag{25}$$

where $(y_l^{l+k}, 0)$ denotes the sequence $\{y_l, y_{l+1}, \ldots, y_{l+k}, 0\}$

Sequence $\{G_{2,n}^*\}$

$$G_{2,n}^*(y_1^{n-1}) = G_{2,n}^{*m}(y_1^{n-1}) = g_n^*(y_1^{n-1}), \quad \text{for } 2 \le n \le m+1$$

$$G^{*}_{2,n}(y_1^{n-1}) = G^{*m}_{2,n}(y_1^{n-1}) = \sum_{k=1}^{t(0,n)} a_k^{(n-1)} \cdot \left[\frac{g_{m+1}{}^{*}(\underline{0},y_1^k,\underline{0}) - g^{*}_{m+1}(\underline{0},y_1^{k-1},\underline{0})}{a^{(m)}_{k-t(-1,n)}} \right]$$

$$+ \sum_{i=0}^{[\frac{n-1}{m}]-1} \sum_{k=t(i,n)+1}^{t(i+1,n)} a_k^{(n-1)} \cdot \left[\frac{g^{*}_{m+1}(y^{k}_{t(i,n)+1},\underline{0}) - g^{*}_{m+1}(y^{k-1}_{t(i,n)+1},\underline{0})}{a^{(m)}_{k-t(i,n)}} \right], \text{ for } n>m+1 \tag{26}$$

where,

$$t(i,n) \stackrel{\Delta}{=} im + n-1 - [\frac{n-1}{m}]\cdot m \tag{27.a}$$

$$(\underline{0},y_1^j,\underline{0}) = \begin{cases} [\underbrace{0,0,0,...0}_{m-t(0,n)},y_1,y_2,...,y_j,\underbrace{0,0,...0}_{t(0,n)-j}] & 1\leq j\leq t(0,n) \\ [\underbrace{0,0,0,...0,0,0,...,0,0,0,...,0}_{m}] & j=0 \end{cases} \tag{27.b}$$

and

$$(y_l^j,\underline{0}) = \begin{cases} [y_l,y_{l+1},...,y_j,\underbrace{0,0,...,0}_{m-(j-l+1)}] & l<j\leq l+m-1 \\ [\underbrace{0,0,...,0,0,0,...,0}_{m}] & j<l \end{cases} \tag{27.c}$$

If the denominator of $\{\cdot\}$ of any term of (25) or if the denominator of $[\cdot]$ of any term of (26) is zero, then that term is not included in the sum.

We observe that the sequences $\{G^{*}_{1,n}\}$ of (25) and $\{G^{*}_{2,n}\}$ of (26) degenerate to the sequence of the optimal at the nominal linear predictors,when in the model in (14), ε=0, (design in the absence of outliers). In addition, using a similar proof as in [18], we can easily show that the sequences $\{G^{*}_{1,n}\}$ and $\{G^{*}_{2,n}\}$ are qualitatively robust, (satisfying condition (B) in Section II), if:

$$\sup_k \sum_{j=1}^{k} |a_j^{(k)}| = c^* < \infty \tag{28}$$

The sequences $\{G^*_{1,n}\}$ and $\{G^*_{2,n}\}$ are identical for m=1. For m>1, these sequences differ: For n > m+1, the predictor $G^*_{1,n}$ is defined in terms of the overlapping sliding blocks of length m observations, while the predictor $G^*_{2,n}$ is defined in terms of disjoint blocks of length m observations.

Asymptotic Performance at the Nominal Process

In this part, we focus on the asymptotic mean squared error induced by the sequences $\{G^*_{1,n}\}$ and $\{G^*_{2,n}\}$ at the nominal process μ_o. We will first assume that μ_o is a zero-mean, stationary Gaussian process and evaluate $e(\mu_o, G_i^{*m})$ where

$$e(\mu_o, G_i^{*m}) = e(\mu_o, G_i^*) = \limsup_{n\to\infty} e_n(\mu_o, G_{i,n}^{*m}), \qquad i=1,2 \tag{29}$$

Then, for a general class of stationary processes, we will obtain an upper bound of (29), which will be tight for small ε.

Fix any $m \geq 1$. Given the infinite past, let the nominal, linear, optimal one-step predictor be:

$$m_o^*(y_{-\infty}^{-1}) = \sum_{i=-\infty}^{-1} d_i y_i \tag{30}$$

If μ_o is also Gaussian, then (30) represents the nominal optimal one-step predictor, given the infinite past. Define:

$$p_m[u_1^m] \stackrel{\Delta}{=} g^*_{m+1}[u_1^m] - g^*_{m+1}[(u_1^{m-1}, 0)] \tag{31}$$

$$q_{m,k}[u_1^m] \stackrel{\Delta}{=} g^*_{m+1}[(u_1^k, \underline{0})] - g^*_{m+1}[(u_1^{k-1}, \underline{0})], \quad 1 \leq k \leq m \tag{32}$$

where g^*_{m+1} is given by (20), $u_1^m = (u_1, u_2, \ldots, u_m) \varepsilon R^m$ and $(u_1^j, \underline{0})$ is given by (27.c).

Then, given the infinite past, the designed robust predictors in (25) and (26) respectively are:

$$G_1^{*m}(y_{-\infty}^{-1}) = G_1^*(y_{-\infty}^{-1}) = \sum_{i=-\infty}^{-1} \frac{d_i}{a_m^{(m)}} \{g^*_{m+1}[y_{i-m+1}^{i}] - g^*_{m+1}[(y_{i-m+1}^{i-1}, 0)]\}$$

$$= \sum_{i=-\infty}^{-1} \frac{d_i}{a_m^{(m)}} \cdot p_m[y_{i-m+1}^{i}] \tag{33}$$

$$G_2^{*m}(y_{-\infty}^{-1}) = G_2^*(y_{-\infty}^{-1}) =$$

$$= \sum_{i=-\infty}^{-1} \sum_{k=1}^{m} \frac{d_{im+k-1}}{a_k^{(m)}} \{g^*_{m+1}[(y_{im}^{im+k-1}, \underline{0}) - g^*_{m+1}[(y_{im}^{im+k-2}, \underline{0})]\}$$

$$= \sum_{i=-\infty}^{-1} \sum_{k=1}^{m} \frac{d_{im+k-1}}{a_k^{(m)}} q_{m,k}[y_{im}^{(i+1)m-1}] \tag{34}$$

where $a_k^{(m)}$ are given as in (25), (26) and $(y_l^j, \underline{0})$ is defined by (27.c).

Let μ_{p_m} be the process induced by μ_o and p_m (31). Then, μ_{p_m} is a zero-mean stationary process. Let $\{R_{p_m}(t)\}_{t=-\infty}^{\infty}$ be the autocorrelations associated with the process μ_{p_m}. That is,

$$R_{p_m}(j-i) = E[p_m(X_{i-m+1}^{i}) p_m(X_{j-m+1}^{j})], \quad -\infty<i,j<\infty \tag{35}$$

Also let,

$$R_{xp_m}(i) = E[X_0 \cdot P_m(X_{i-m+1}^{i})], \qquad i \le -1 \tag{36}$$

Similarly, we define,

$$R_{q_{m,k,l}}(j-i) = E[q_{m,k}(X_{im}^{(i+1)m-1}) \cdot q_{m,l}(X_{jm}^{(j+1)m-1})], \qquad \begin{matrix} 1 \le k, l \le m \\ -\infty < i,j < \infty \end{matrix} \tag{37}$$

$$R_{xq_{m,k}}(i) = E[X_o q_{m,k}(X_{im}^{(i+1)m-1})], \qquad \begin{matrix} i \le -1 \\ 1 \le k \le m \end{matrix} \tag{38}$$

We will express $c(\mu_o, G_i^{*m})$ in (29), in terms of the quantities in (35)-(38). These quantities are non-trivial to obtain, since the mappings p_m and q_m are nonlinear. We will determine these quantities assuming that μ_o is Gaussian.

Assume that μ_o is a Gaussian source. Let,

$$Z_{1,i} = m_o(X_{i-m+1}^{i}) = B_m^T X_{i-m+1}^{i}$$

$$= \sum_{t=1}^{m} a_t^{(m)} X_{i-m+t} \tag{39}$$

$$Z_{2,i} = \sum_{t=1}^{m-1} a_t^{(m)} X_{i-m+t} \tag{40}$$

Then, $\{Z_{1,i}\}$ and $\{Z_{2i}\}$ are zero-mean stationary, Gaussian processes. Let $\{R_{11}(j)\}$ and $\{R_{22}(j)\}$ be the autocorrelations associated with these processes respectively, and let $\{\rho_{11}(j)\}$, $\{\rho_{22}(j)\}$ be the associated correlation coefficients. Then,

$$R_{11}(j-i) = E[Z_{1,i} Z_{1,j}] \overset{\Delta}{=} \rho_{11}(j-i) R_{11}(0) \qquad -\infty < i,j < \infty \tag{41}$$

$$R_{22}(j-i) \overset{\Delta}{=} E[Z_{2,i} Z_{2,j}] \overset{\Delta}{=} \rho_{22}(j-i) R_{22}(0) \qquad -\infty < i,j < \infty \tag{42}$$

Let,

$$R_{12}(j-i) = E[Z_{1,i}Z_{2,j}] \overset{\Delta}{=} \rho_{12}(j-i)\sqrt{R_{11}(0)R_{22}(0)}, \quad -\infty<i,j<\infty \tag{43a}$$

$$R_{21}(j-i) = E[Z_{2,i}Z_{1,j}] \overset{\Delta}{=} \rho_{21}(j-i)\sqrt{R_{11}(0)R_{22}(0)}, \quad -\infty<i,j<\infty \tag{43b}$$

Also, let $R_{\infty}(0) = \sigma^2 = E[X_i^2]$ and

$$R_{0t}(i) = E[X_0 Z_{t,i}] \overset{\Delta}{=} \rho_{0t}(i)\sigma\sqrt{R_{tt}(0)}, \quad t=1,2;\ i\leq -1 \tag{43c}$$

Define:

$$W_{k,i} = \sum_{t=1}^{k} a_t^{(m)} X_{im+t-1}, \quad 0\leq k\leq m,\ i\leq -1 \tag{44}$$

Then, $W_{0,i} = 0$, and for each $k\geq 1$, $\{W_{k,i}\}$ is a zero-mean, stationary Gaussian process. Let,

$$R_{w_{k,l}}(j-i) = E[W_{k,i}W_{l,j}], \quad \begin{matrix} k,l = 0,1,\dots,m \\ -\infty<i,j<\infty \end{matrix} \tag{45}$$

and let,

$$\rho_{w_{k,l}}(j-i) = \begin{cases} \dfrac{R_{w_{k,l}}(j-i)}{(R_{w_{k,k}}(0)R_{w_{l,l}}(0))^{1/2}} & \text{if exists} \\ 0, & \text{otherwise} \end{cases} \tag{46}$$

Let

$$R_{xw_k}(i) = E[X,W_{k,i}], \quad 0\leq k\leq m;\ i\leq -1 \tag{47}$$

$$\rho_{xw_k}(i) = \begin{cases} \dfrac{R_{xw_k}(i)}{[\sigma^2 \cdot R_{w_{k,k}}(0)]^{1/2}}, & \text{if exists} \\ 0 & \text{otherwise} \end{cases} \tag{48}$$

Let $\phi(x)$ and $\Phi(x)$ respectively be the standard normal density zero mean, unit variance, and its cumulative distribution function, evaluated at the point x. Also, let $\phi^{(n)}(u)$ be the nth derivative of $\phi(x)$ w.r.t. x, evaluated at u. Fix any $\lambda>0$. Let,

$$h_\lambda(u) = \begin{cases} u, & |u| \leq \lambda \\ \lambda \operatorname{sgn}(u), & \text{otherwise} \end{cases} \tag{49}$$

For any σ_1, σ_2 and ρ such that $\sigma_1>0$, $\sigma_2 \geq 0$ and $|\rho|<1$, let us define:

$$\theta[\lambda,\sigma_1] \stackrel{\Delta}{=} \sigma_1^2 \cdot \Phi(\frac{\lambda}{\sigma_1}) + (2\lambda^2 - \sigma_1^2)\Phi\left[\frac{-\lambda}{\sigma_1}\right] \tag{50}$$

$$\begin{aligned} \beta[\lambda,\sigma_1,\sigma_2,\rho] &\stackrel{\Delta}{=} A[\lambda,\sigma_1,\sigma_2,\rho] \cdot I[\lambda,\sigma_1,1] + \\ &+ [\sigma_2^2(1-\rho^2)]^{1/2} \sum_{n=0}^{\infty} F[\sigma_1,\rho,n] \cdot \beta[\lambda,\sigma_2,\rho,n] \cdot I[\lambda,\sigma_1,n] \\ &+ \lambda \sum_{n=1}^{\infty} F[\sigma_1,\rho,n] \cdot \beta[\lambda,\sigma_2,\rho,n-1] \cdot I[\lambda,\sigma_1,n] \\ &+ 2[\sigma_2^2(1-\rho^2)]^{1/2} \sum_{n=2}^{\infty} F[\sigma_1,\rho,n] \cdot \beta[\lambda,\sigma_2,\rho,n-2] \cdot I[\lambda,\sigma_1,n] \end{aligned} \tag{51}$$

where,

$$A[\lambda,\sigma_1,\sigma_2,\rho] = \gamma_c\cdot[\Phi\left(\frac{\lambda}{\sigma_c}\right) - \Phi\left(\frac{-\lambda}{\sigma_c}\right)] \tag{52}$$

$$\beta[\lambda,\sigma_2,\rho,n] = \phi^{(n)}\left(\frac{\lambda}{\sigma_c}\right) - \phi^{(n)}\left(\frac{-\lambda}{\sigma_c}\right) \tag{53}$$

$$F[\sigma_1,\rho,n] = \left(\frac{\rho}{\sqrt{1-\rho^2}}\right)^n \cdot\frac{1}{\sigma_1^n}\cdot\frac{1}{n!} \tag{54}$$

$$I[\lambda,\sigma_1,n] = \int_{-\infty}^{\infty} h_\lambda(x)\cdot x^n\cdot\frac{\phi\left(\frac{x}{\sigma_1}\right)}{\sigma_1}\,dx \tag{55}$$

and where

$$\sigma_c^2 = \sigma_2^2(1-\rho^2) \tag{56a}$$

$$\gamma_c = \rho\,\frac{\sigma_2}{\sigma_1} \tag{56b}$$

Also, let $\zeta[\lambda,\sigma_1,\sigma_2,\rho]$ be given by the r.h.s. of (51) except that we replace $I[\lambda,\sigma_1,1]$ with $I[\infty,\sigma_1,1]$, and $I[\lambda,\sigma_1,n]$ with $I[\infty,\sigma_1,n]$. For $\sigma_1 = \sigma_2$ and $\rho=1$, we define:

$$\beta[\lambda,\sigma_1,\sigma_1,1] = \theta[\lambda,\sigma_1] \tag{57}$$

Notice that the computation of θ does not involve any series where as for $|\rho|<1$, β and ζ are given in terms of series. The definition (52) is also consistent with the meaning associated with θ and β, (see (B.2) and (B.3) of the proof of Theorem 2.) Using direct but tedious computations, we obtain:

$$\beta[\lambda,\sigma_2,\rho,n] = \begin{cases} \left(\frac{2}{\pi}\right)^{1/2} \exp(-\lambda^2\sigma_c^{-2})\sum_{t=0}^{k}(-1)^{t+1}\left(\frac{\lambda}{\sigma_c}\right)^{2k-2t+1}\frac{(2k+1)!2^{-t}}{(2k-2t+1)!t!}, & \text{if n odd, } n=2k+1 \\ 0 \quad , & \text{if n even} \end{cases} \tag{58}$$

Also, if n is even, then $I[\lambda,\sigma_1,n] = I[\infty,\sigma_1,n] = 0$. If $n=2k+1$, then:

$$I[\lambda,\sigma_1,2k+1] = \frac{1}{\sqrt{\pi}}\left(\frac{\lambda}{\alpha}\right)^{2k+2}\left\{\sqrt{\pi}[\Phi(\frac{\lambda}{\sigma_1})-\Phi(-\frac{\lambda}{\sigma_1})]\cdot\prod_{t=0}^{k}(k-t+\frac{1}{2})\right.$$

$$\left. -e^{-\alpha^2}\cdot\alpha^{2k+1}[1+\sum_{l=1}^{k}\alpha^{-2l}\cdot\prod_{t=0}^{l-1}(k-t+\frac{1}{2})]\right\}$$

$$+\frac{1}{\sqrt{\pi}}\cdot 2^{k+1}\cdot\sigma_1^{2k+2}\cdot k!\cdot e^{-\alpha^2}\cdot\sum_{j=0}^{k}\frac{\alpha^{j+1}}{j!} \tag{59}$$

$$I[\infty,\sigma_1,2k+1] = \sigma_1^{2k+2}\cdot\prod_{t=1}^{k+1}(2t-1) \tag{60}$$

where

$$\alpha = \frac{\lambda}{\sqrt{2}\sigma_1} \tag{61}$$

We are now ready to obtain $e(\mu_0,G_i^{*m})$ in (29), assuming certain conditions. Let N be the set of nonnegative integers. Let J be the Borel σ-field generated by the discrete topology on N. Let (NxR, $JX\mathcal{B}$) be the product space where $JX\mathcal{B}$ is the Borel σ-field generated by the product topology on NXR. Let ν be the product measure on $JX\mathcal{B}$, product of the counting measure on J and Lebesque measure on $\mathcal{B}$. Define,

$$r(n,x\mid\lambda,\sigma_1,\sigma_2,\rho) \overset{\Delta}{=} \frac{\phi^{(n)}(\frac{\lambda}{\sigma_c})\cdot(\frac{\gamma_c}{\sigma_c})^n}{n!}\cdot x^{n+1}\cdot\frac{\phi(\frac{x}{\sigma_1})}{\sigma_1} \tag{62}$$

where σ_c and γ_c are given by (56.a) and (56.b) respectively.

Theorem 2: Let μ_o be a zero-mean, stationary Gaussian source with variance σ^2. Assume that (28) holds. Then,

$$c(\mu_o,G_1^{*m}) = c(\mu_o,G_1^*) = \sigma^2 - 2\sum_{i=-\infty}^{-1}\frac{d_i}{a_m^{(m)}}R_{xp_m}(i) + \sum_{i=-\infty}^{-1}\sum_{j=-\infty}^{-1}\frac{d_i d_j}{[a_m^{(m)}]^2}R_{p_m}(j-i) \tag{63}$$

$$c(\mu_o,G_2^{*m}) = c(\mu_o,G_2^*) = \sigma^2 - 2\sum_{i=-\infty}^{-1}\sum_{k=1}^{m}\frac{d_{im+k-1}}{a_k^{(m)}}R_{xq_{m,k}}(i)$$

$$+\sum_{i=-\infty}^{-1}\sum_{j=-\infty}^{-1}\sum_{k=1}^{m}\sum_{l=1}^{m}\frac{d_{im+k-1}}{a_k^{(m)}}\cdot\frac{d_{jm+l-1}}{a_l^{(m)}}\cdot R_{q_{m,k,l}}(j-i) \tag{64}$$

where $\{d_i\}$, $a_k^{(m)}$ are given as in (33) and (34), and R_{p_m}, R_{xp_m}, $R_{q_{m,k,l}}$ and $R_{xq_{m,k}}$ are given by (35)-(38). If

$$\int r(n,x\mid\lambda,\sigma_1,\sigma_2,\rho)dv \tag{65}$$

exists (i.e. the integral is not of the form $\infty,-\infty$ in the sense of Lebesque) for all tuples $(\lambda,\sigma_1,\sigma_2,\rho)$ which are the arguments of β and ζ of (66) and (67) below, then,.

$$R_{p_m}(j-i) = \sum_{l,t=1}^{2}(-1)^{l+t}\beta[\lambda_m, \sqrt{R_{ll}(0)}, \sqrt{R_{tt}(0)}, \rho_{lt}(j-i)], \quad -\infty<i,j<\infty \tag{66}$$

$$R_{xp_m}(i) = \zeta[\lambda_m, \sigma, \sqrt{R_{11}(0)}, \rho_{01}(i)] - \zeta[\lambda_m, \sigma, \sqrt{R_{22}(0)}, \rho_{02}(i)] \qquad i \leq -1 \tag{67}$$

If (65) exists for all tuples $(\lambda, \sigma_1, \sigma_2, \rho)$ which are the arguments of β and ζ of (68) and (69) below, then,

$$R_{q_{m,k,l}}(j-i) =$$

$$= \sum_{s,t=0}^{1} (-1)^{s+t} \beta[\lambda_m, \sqrt{R_{w_{k-s,k-s}}(0)}, \sqrt{R_{w_{l-t,l-t}}(0)}, \rho_{w_{k-s,l-t}}(j-i)] \tag{68}$$

$$, 1 \leq k, l \leq m$$

$$-\infty < i, j < \infty$$

$$R_{xq_{m,k}}(i) = \zeta[\lambda_m, \sigma, \sqrt{R_{w_{k,k}}(0)}, \rho_{xw_k}(i)] - \zeta[\lambda_m, \sigma, \sqrt{R_{w_{k-1,k-1}}(0)}, \rho^{(i)}_{xw_{k-1}}] \tag{69}$$

$$1 \leq k \leq m$$

$$i \leq -1$$

Proof: See the Appendix.

Remarks:

(1) The integral in (65) involves four parameters, $\lambda, \sigma_1, \sigma_2$ and ρ, and its existence is required to ensure that Fubini's theorem is applicable. By Corollary 2.65 of Ash [21], the integral will exist if

$$\sum_{n=1}^{\infty} \int_{-\infty}^{\infty} | r(n, x | \lambda, \sigma_1, \sigma_2, \rho) | dx < \infty \tag{70}$$

Since,

$$\sup_x |\phi^{(n)}(x)| < \frac{2^{\frac{n+1}{2}} \cdot (\frac{n+1}{2})!}{\sqrt{2\pi e^n}}, \quad n \geq 1, \text{ n odd} \tag{71}$$

and since,

$$\int_{-\infty}^{\infty} x^{n+1} \cdot \frac{\phi(\frac{x}{\sigma_1})}{\sigma_1} dx = \sigma_1^{n+1} \frac{n!}{2^{\frac{n-1}{2}} \cdot (\frac{n-1}{2})!}, \quad n \geq 1, \text{ n odd} \tag{72}$$

hence,

$$\sum_{n=1}^{\infty} \int_{-\infty}^{\infty} |r(n,x|\lambda,\sigma_1,\sigma_2,\rho)| dx \leq \sum_{n=1}^{\infty} \frac{\sigma_1}{\sqrt{2\pi}} \left(\frac{\rho}{\sqrt{e} \cdot \sqrt{1-\rho^2}}\right)^n (n+1), \quad \text{n odd} \tag{73}$$

The series in the upper bound of (73) converges if,

$$\rho < \left[\frac{e}{e+1}\right]^{1/2} \tag{74}$$

Similarly, summing over even values of n of the series in (70), we arrive at (74). Hence, (74) is a sufficient condition for the existence of the integral in (65). In this case, $c(\mu_o, G_i^{*m})$ can be approximated by considering only a finite number of terms of all the series of β and ζ.

(2) Hölder inequality gives a simple condition for checking whether the series in the definition of β or ζ is divergent. Define N_1, N_2 as in Lemma 1 of the Appendix. Then by (B.3) and Hölder,

$$|\beta(\lambda,\sigma_1,\sigma_2,\rho)| = |E[h_\lambda(N_1)h_\lambda(N_2)]|$$

$$\leq \{E[h_\lambda(N_1)]^2\}^{1/2}\cdot\{E[h_\lambda(N_2)]^2\}^{1/2}$$

$$= \{\theta[\lambda,\sigma_2]\}^{1/2}\cdot\{\theta[\lambda,\sigma_2]\}^{1/2} \tag{75}$$

The upper bound of (75) does not involve any series and can be easily computed using (50). If the series of β diverges, then (75) will not hold. Similarly,

$$|\zeta[\lambda,\sigma_1,\sigma_2,\rho]| \leq \sigma_1\cdot\{\theta[\lambda,\sigma_2]\}^{1/2} \tag{76}$$

We will now obtain an upper bound if $e(\mu_o\cdot G_1^{*m})$, which is also applicable to non-Gaussian processes, which does not require any restriction as that in (65), which is easy to compute, and which is directly related to $e(\mu_o,m_o^*)$, where $e(\mu_o,m_o^*)$ is the asymptotic mean squared error at μ_o that is induced by the nominal optimal linear predictors $m_o(y_1^{n-1})$. Using the proof of Theorem 3 below an upper bound of $e(\mu_o,G_2^{*m})$ can also be obtained in a similar manner. Let,

$$H_m = E[(|X_m| + \frac{2\lambda_m}{|a_m^{(m)}|})^2\{1-1_{J^m}(X^m)\}] \tag{77}$$

where

$$J^m = \{y^m : |\sum_{j=1}^{m} a_j^{(m)} y_j| < \lambda_m \text{ and } |\sum_{j=1}^{m-1} a_j^{(m)} y_i| < \lambda_m\} \tag{78}$$

Also, let

$$D^* = \sum_{i=-\infty}^{-1} |d_i| \tag{79}$$

Theorem 3: Let $\{X_k\}_{k=-\infty}^{\infty}$ be a zero-mean, finite variance σ^2, stationary source with distribution

μ_o. Fix any $m \geq 1$. Then,

$$\text{(a)} \qquad e(\mu_o, G_1^{*m}) = e(\mu_o, G_1^*) \leq (\{e(\mu_o, m_o^*)\}^{1/2} + D^* \sqrt{H_m})^2 \tag{80}$$

$$\text{(b)} \qquad \lim_{\varepsilon \to 0} (\{e(\mu_o, m_o^*)\}^{1/2} + D^* \sqrt{H_m})^2 = e(\mu_o, m_o^*) \tag{81}$$

Proof: See the Appendix.

Remark: For Gaussian sources, if m=1, then,

$$H_1 = 2 \int_{\frac{\lambda_1}{|a_1^{(1)}|}}^{\infty} (u + \frac{2\lambda_1}{|a_1^{(1)}|})^2 \cdot \frac{\phi(\frac{u}{\sigma})}{\sigma} du \tag{82}$$

For m>1, we obtain an upper bound of H_m which is easy to compute, as follows. By (C.7) in the Apppendix, we have:

$$E[X_m^2(1-1_{J_t^m}(X^m))] \leq \{E[X_m^4]\}^{1/2} \cdot \{E[1-1_{J_t^m}(X^m)]\}^{1/2} =$$

$$= \sqrt{3}\sigma^2 \cdot \{E[1-1_{J_t^m}(X^m)]\}^{1/2}, \qquad t=1,2$$

$$E[|X_m|(1-1_{J_t^m}(X^m))] \leq \sigma\{E[1-1_{J_t^m}(X^m)]\}^{1/2}, \qquad t=1,2$$

Therefore, by (77) and (C.6) in the Appendix, we obtain:

$$H_m \leq \sqrt{2}\left\{\sqrt{3}\sigma^2 + \frac{4\lambda_m \sigma}{|a_m^{(m)}|} + \frac{4\lambda_m^2}{|a_m^{(m)}|^2}\right\}^{1/2} \cdot \left[\left\{1-\Phi(\frac{\lambda_m}{\sigma_{1,m}})\right\}^{1/2} + \left\{1-\Phi\left[\frac{\lambda_m}{\sigma_{2,m}}\right]\right\}^{1/2}\right] \tag{83}$$

where $\sigma_{t,m}^2$ is the variance of $\sum_{j=1}^{m-t+1} a_j^{(m)} X_j$, j=1,2.

IV. BREAKDOWN POINT AND INFLUENCE FUNCTION

In this section, we obtain the breakdown points and the influence functions of the prediction operations in (29) (m=1 only) and the operations in (30), following the definitions (9)-(16). We first consider the case of the per-datum outlier model in (6) and the operations in (29) and (30), for m=1. Then, we consider the case of m>1. It is easy to verify that the breakdown point of the nominal optimal linear predictors in (24) is zero, for any zero-mean, finite variance σ^2, stationary process, μ_o. Let $I_{0,1}(w)$ be the influence function of these nominal predictors for the contamination model in (6). Then,

$$I_{0,1}(w) = (w^2-\sigma^2)\cdot \sum_{i=-\infty}^{-1} d_i^2 - 2 \sum_{i=-\infty}^{-1} \sum_{j=-\infty}^{-1} d_i d_j R_x(j-i) + 2 \sum_{i=-\infty}^{-1} d_i R_x(i) \tag{84}$$

where $\{d_i\}_{i=-\infty}^{-1}$ is given by (30) and $\{R_x(i)\}_{i=-\infty}^{\infty}$ is the autocorrelation function of the process μ_o.

Consider now the predictors in (29) and (30), which are designed assuming that the nominal process is μ_o and the level of contamination is ε. Asymptotically, these predictors are given by (33) and (34). Let m=1. Then (33) and (34) coincide. Let G* denote the asymptotic predictor in (33) (or (34)) for the case of m=1. Fix any w and let the contaminating process be deterministic with amplitude w. Let the level of contamination be δ. Then,

$$e(\mu_{\delta,w},G^*) = e(\mu_{\delta,w,1},G_1^{*1}) =$$

$$=E[\{X_0-G^*(Y_{-\infty}^{-1})\}^2|\delta,w]=E[\{X_0-\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}p_1(Y_i)\}^2|\delta,w]$$

$$=\sigma^2-2\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}\cdot E[X_0\cdot p_1(Y_i)|\delta,w]+\sum_{i=-\infty}^{-1}\sum_{j=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}\cdot\frac{d_j}{a_1^{(1)}}\cdot E[p_1(Y_i)p_1(Y_j)|\delta,w] \tag{85}$$

Now, for any $i\leq-1$ and by (6), we have,

$$E[X_0\cdot p_1(Y_i)|\delta,w]=(1-\delta)E[X_0\cdot p_1(X_i)]+\delta E[X_0\cdot p_1(w)] \tag{86}$$
$$=(1-\delta)R_{xp_1}(i)$$

where $R_{xp_1}(i)$ is defined by (36). Also, for $i=j\leq-1$, we have,

$$E[p_1(Y_i)p_1(Y_j)|\delta,w]=E[\{p_1(Y_i)\}^2|\delta,w]=(1-\delta)E[p_1(X_i)]^2+\delta E[p_1(w)]^2$$

$$=(1-\delta)R_{p_1}(0)+\delta[p_1(w)]^2 \tag{87}$$

and for $i\neq j$

$$E[p_1(Y_i)p_1(Y_j)|\delta,w]=(1-\delta)^2E[p_1(X_i)p_1(X_j)]+\delta^2[p_1(w)]^2$$

$$+\delta(1-\delta)\{E[p_1(X_i)p_1(w)]+E[p_1(X_j)p_1(w)]\}$$

$$=(1-\delta)^2\cdot R_{p_1}(j-i)+\delta^2[p_1(w)]^2 \tag{88}$$

where $\{R_{p_1}(j-i)\}$ is defined by (35).

Using (86)-(87), we can determine the breakdown point in (11) of the predictors in (29) (or (30)), for m=1, as:

$$\varepsilon^*_{G_1^{*1}}=\varepsilon^*_{G^*}=\sup_{0\leq\delta<1}\{\delta:\ \sup_w e(\mu_{\delta,w},G_1^{*1})\leq\sigma^2\} \tag{89}$$

Equivalently,

$$\varepsilon^*_{G*} = \sup_{0<\delta\le 1}\left\{ \delta:\ \sum_{i=-\infty}^{-1}\left[\frac{d_i}{a_1^{(1)}}\right]^2 \left\{(1-\delta)R_{p_1}(0)+\delta\lambda_1^2\right\}\right.$$

$$\left. + \sum_{i=-\infty}^{-1}\sum_{j=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}\cdot\frac{d_j}{a_1^{(1)}}\cdot\{(1-\delta)^2R_{p_1}(j-i)+\delta^2\cdot\lambda_1^2\} \le 2(1-\delta)\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}R_{xp_1}(i)\right\} \tag{90}$$

Notice that for m=1, $\varepsilon^*_{G*} = \bar{\varepsilon}^*_{G*}$ ((12)).

We now determine the influence function substituting (86)-(88) in (85), we obtain:

$$e(\mu_{\delta,w},G^*) - e(\mu_o,G^*) =$$

$$= E[\{X_0-G^*(Y_{-\infty}^{-1})\}^2 \mid \delta,w] - E[X_0-G^*(X_{\infty}^{-1})]^2 =$$

$$= -2\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}\cdot\{E[X_0p_1(Y_i)\mid\delta,w] - E[X_0\cdot p_1(X_i)]\}$$

$$+ \sum_{i=-\infty}^{-1}\sum_{j=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}\cdot\frac{d_j}{a_1^{(1)}}\cdot\{E[p_1(Y_i)p_1(Y_j)\mid\delta,w] - E[p_1(X_i)p_1(X_j)]\}$$

$$= 2\delta\cdot\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}\cdot R_{xp_1}(i) + \delta\sum_{i=-\infty}^{-1}\left(\frac{d_i}{a_1^{(1)}}\right)^2\cdot\{[p_1(w)]^2 - R_{p_1}(0)\}$$

$$+ \sum_{i=-\infty}^{-1}\sum_{j=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}\cdot\frac{d_j}{a_1^{(1)}}\cdot\{\delta^2[p_1(w)]^2 + (-2\delta+\delta^2)R_{p_1}(j-i)\} \tag{91}$$

Therefore, the influence function $I_{G*}(w)$ for the designed predictors in (29) (or (30)), for m=1, is:

$$I_{G^*}(w) = 2\sum_{i=-\infty}^{-1} \frac{d_i}{a_1^{(1)}} R_{xp_1}(i) + \sum_{i=-\infty}^{-1} \left(\frac{d_i}{a_1^{(1)}}\right)^2 \cdot \{[p_1(w)]^2 - R_{p_1}(0)\}$$

$$-2\sum_{i=-\infty}^{-1}\sum_{j=-\infty}^{-1} \frac{d_i}{a_1^{(1)}} \cdot \frac{d_j}{a_1^{(1)}} \cdot R_{p_1}(j-i) \tag{92}$$

Notice that for $\varepsilon>0$, $I_{G*}(w)$ is a bounded and continuous function of w, in contrast to $I_{0,1}(w)$, in (84). Also, if $\varepsilon=0$, then (92) reduces to (84).

The above expressions, (90) and (92), of the breakdown point and the influence function respectively; require the knowledge of $\{R_{p_1}(i)\}_{i=-\infty}^{\infty}$ and $\{R_{xp_1}(i)\}_{i=-\infty}^{-1}$. If μ_o is also assumed to be Gaussian, then we can determine those quantities using Theorem 2, if (65) holds. For non-Gaussian sources, or for Gaussian sources where (65) does not hold, it may not be possible to determine $\{R_{p_1}(i)\}$ and $\{R_{xp_1}(i)\}$ analytically.

We now take an alternate approach and determine an upper bound of the influence function and a lower bound of the breakdown point, for nominal processes that are not necessarily Gaussian. We have:

$$e(\mu_{\delta,w}, G^*) = E[\{X_0 - G^*(Y_{-\infty}^{-1})\}^2 \mid \delta, w] = E[X_0 - \sum_{i=-\infty}^{-1} \frac{d_i}{a_1^{(1)}} p_1(X_i)]^2$$

$$+ 2E\{[X_0 - \sum_{i=-\infty}^{-1} \frac{d_i}{a_1^{(1)}} p_1(X_i)] \cdot [\sum_{i=-\infty}^{-1} \frac{d_i}{a_1^{(1)}} (p_1(X_i) - p_1(Y_i))] \mid \delta, w\}$$

$$+ E[\{\sum_{i=-\infty}^{-1} \frac{d_i}{a_1^{(1)}} (p_1(X_i) - p_1(Y_i))\}^2 \mid \delta, w] \tag{93}$$

Now,

$$E\{[X_0-\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}p_1(X_i)]\cdot[\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}(p_1(X_i)-p_1(Y_i))]\mid\delta,w\}$$

$$=\delta\cdot E\{[X_0-\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}p_1(X_i)]\cdot[\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}(p_1(X_i)-p_1(w))]\}$$

$$=\delta\cdot E\{[X_0-\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}p_1(X_i)]\cdot[\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}p_1(X_i)]\}$$

$$\leq\delta\{E[X_0-\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}p_1(X_i)]^2\}^{1/2}\cdot\{E[\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}p_1(X_i)]^2\}^{1/2}$$

$$\leq\delta\cdot\{E[X_0-\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}p_1(X_i)]^2\}^{1/2}\cdot[\sum_{i=-\infty}^{-1}\left|\frac{d_i}{a_1^{(1)}}\right|\cdot R_{p_1}(0)] \tag{94}$$

The last two inequalities of (94) follow respectively by Holder and Minkowaski. Also, by Minkowaski we have:

$$\{E[\{\sum_{i=-\infty}^{-1}\frac{d_i}{a_1^{(1)}}(p_1(X_i)-p_1(Y_i))\}^2\mid\delta,w]\}^{1/2}$$

$$\leq\sum_{i=-\infty}^{-1}\left|\frac{d_i}{a_1^{(1)}}\right|\cdot\{E[\{p_1(X_i)-p_1(Y_i)\}^2\mid\delta,w]\}^{1/2}$$

$$=\delta^{1/2}\cdot\sum_{i=-\infty}^{-1}\left|\frac{d_i}{a_1^{(1)}}\right|\cdot\{E[p_1(X_i)-p_1(w)]^2\}^{1/2} \tag{95}$$

Therefore,

$$E[\{\sum_{i=-\infty}^{-1} \frac{d_i}{a_1^{(1)}}(p_1(X_i) - p_1(Y_i))\}^2 \mid \delta, w] \le \delta(\sum_{i=-\infty}^{-1} \left| \frac{d_i}{a_1^{(1)}} \right|)^2 \cdot (R_{p_1}(0) + [p_1(w)]^2) \tag{96}$$

Using (93), (94), (96), and Theorem 3, a lower bound ε_{G*}^{*l} of the breakdown point ε_{G*}^{*} is obtained as:

$$\varepsilon_{G*}^{*l} = \sup_{\delta} \{\delta : (\sqrt{c(\mu_o, m_o^*)} + D^*\sqrt{H_1})^2 + 2\cdot\delta\cdot[\sum_{i=-\infty}^{-1} \left| \frac{d_i}{a_1^{(1)}} \right| \cdot R_{p_1}(0)]\cdot(\sqrt{c(\mu_o, m_o^*)} + D^*\sqrt{H_1}),$$

$$0 \le \delta < 1$$

$$+ \delta\cdot(\sum_{i=-\infty}^{-1} \left| \frac{d_i}{a_1^{(1)}} \right|)^2 \cdot (R_{p_1}(0) + \lambda_1^2) \le \sigma^2\} \tag{97}$$

For small ε, the lower bound ε_{G*}^{*l} is strictly positive by Theorem 3b. Also, an upper bound $I_{G*}^{u}[w]$ of the influence function $I_{G*}[w]$ in (92) is obtained, using (93), (94), (96), and (80). The bound is as follows:

$$I_{G*}^{u}[w] = 2(\sqrt{c(\mu_o, m_o^*)} + D^*\sqrt{H_1})\cdot[\sum_{i=-\infty}^{-1} \left| \frac{d_i}{a_1^{(1)}} \right| R_{p_1}(0)] + \left[\sum_{i=-\infty}^{-1} \left| \frac{d_i}{a_1^{(1)}} \right| \right]^2 \cdot (R_{p_1}(0) + [p_1(w)]^2) \tag{98}$$

The upper bound $I_{G*}^{u}[w]$ is a bounded function of w, if $\varepsilon > 0$.

We now consider the case of m > 1, and for the predictors in (30), we determine the breakdown point and the influence function assuming that the observation process corresponds to the m-size (block) outlier model of (14). Fix any w and let the contaminating process be deterministic with amplitude w. Let the level of contamination be δ. Then, by (34) and the model in (14), and following the steps (86)-(88), we obtain:

$$c(\mu_{\delta,w,m},G_2^{*m}) =$$

$$= E[\{X_0 - G_2^{*m}(Y_{-\infty}^{-1})\}^2 \mid \delta, w^m] = \sigma^2 - 2(1-\delta) \sum_{i=-\infty}^{-1} \sum_{k=1}^{m} \frac{d_{im+k-1}}{a_k^{(m)}} R_{xq_{m,k}}(i)$$

$$+ \sum_{i=-\infty}^{-1} \sum_{k=1}^{m} \sum_{l=1}^{m} \frac{d_{im+k-1}}{a_k^{(m)}} \cdot \frac{d_{im+l-1}}{a_l^{(m)}} \{(1-\delta) R_{q_{m,k,l}}(0) + \delta(q_{m,k}[w^m] \cdot q_{m,l}[w^m])\}$$

$$+ \sum_{i=-\infty}^{-1} \sum_{j=-\infty}^{-1} \sum_{k=1}^{m} \sum_{l=1}^{m} \frac{d_{im+k-1}}{a_k^{(m)}} \cdot \frac{d_{jm+l-1}}{a_l^{(m)}} \{(1-\delta)^2 R_{q_{m,k,l}}(j-i) + \delta^2 (q_{m,k}[w^m] \cdot q_{m,l}[w^m])\} \cdot \tag{99}$$

In (99), w^m is the length-m vector with all its elements being w. Also, we have used w^m in the conditioning, to indicate that the contamination model in (14) is used. Using (99), we obtain the breakdown point $\varepsilon_{G_2^{*m},m}$, (15), and the influence function $I_{G_2^{*m},m}[w]$, (16), directly by the definitions below.

$$\varepsilon^*_{G_2^{*m},m} = \sup_{0 \le \delta \le 1} \{\delta : \sup_w c(\mu_{\delta,w,m},G_2^{*m}) \le \sigma^2\} \tag{100}$$

$$I_{G_2^{*m},m}[w] = 2 \sum_{i=-\infty}^{-1} \sum_{k=1}^{m} \frac{d_{im+k-1}}{a_k^{(m)}} R_{xq_{m,k}}(i)$$

$$+ \sum_{i=-\infty}^{-1} \sum_{k=1}^{m} \sum_{l=1}^{m} \frac{d_{im+k-1}}{a_k^{(m)}} \cdot \frac{d_{im+l-1}}{a_l^{(m)}} \{(q_{m,k}[w^m] q_{m,l}[w^m] - R_{q_{m,k,l}}(0)\}$$

$$-2 \sum_{i=-\infty}^{-1} \sum_{j=-\infty}^{-1} \sum_{k=1}^{m} \sum_{l=1}^{m} \frac{d_{im+k-1}}{a_k^{(m)}} \cdot \frac{d_{jm+l-1}}{a_l^{(m)}} \cdot R_{q_{m,k,l}}(j-i) \tag{101}$$

For non-Gaussian sources, an upper bound of the influence function and a lower bound of the breakdown point for m > 1 and for the block contamination model (14) can be obtained, using

Theorem 3 and steps similar to those taken to obtain the results for m=1.

For m > 1, for the per-datum outlier model in (6), and for the predictors in (29) which are designed assuming the level of contaimination is ε, $\varepsilon>0$, we have:

$$e(\mu_{\delta,w,1}, G_1^{*m}) = E[\{X_0 - G_1^*(Y_{-\infty}^{-1})\}^2 \mid \delta, w] =$$

$$= \sigma^2 - 2\sum_{i=-\infty}^{-1} \frac{d_i}{a_m^{(m)}} E[X_0 \cdot p_m(Y_{i-m+1}^{i}) \mid \delta, w] + \sum_{i=-\infty}^{-1}\sum_{j=-\infty}^{-1} \frac{d_i d_j}{[a_m^{(m)}]^2} \cdot E[p_m(Y_{i-m+1}^{i}) p_m(Y_{j-m+1}^{j}) \mid \delta, w] \tag{102}$$

Hence,

$$\sup_w e(\mu_{\delta,w,1}, G_1^{*m}) \tag{103}$$

is continuous in δ at $\delta = 0$. Also, if $\delta=0$, then $e(\mu_{\delta,w,1}, G_1^{*m})$ equals the asymptotic mean squared error $e(\mu_o, G_1^*)$ induced by the predictors $\{G_{1,n}^{*m}\}$ at the nominal process μ_o. Hence, by the definition of the breakdown point in (11), the breakdown point of the designed predictors $\{G_{1,n}^{*m}\}$ that corresponds to the per-datum outlier model in (6), viz. $\varepsilon^*_{G_1^{*m}}$ in (11), will be positive if and only if

$$e(\mu_o, G_1^{*m}) < \sigma^2 \tag{104}$$

Similar conclusions can be drawn for any size (batch) of outliers using the model in (14).

V. NUMERICAL EXAMPLES

Let μ_o be a zero-mean, stationary, Gaussian process. Also, let μ_o be auto-regressive. Consider the following representations of the nominal process μ_o:

Nominal Process 1:

$$x_k = 0.7x_{k-1} + w_k \qquad k=\ldots,-1,0,1,\ldots$$

Nominal Process 2:

$$x_k = 0.7x_{k-1} - 0.3x_{k-2} + w_k \qquad k=\ldots,-1,0,1,\ldots$$

Nominal Process 3:

$$x_k = 0.7x_{k-1} - 0.3x_{k-2} + 0.2x_{k-3} + w_k \qquad k=\ldots,-1,0,1,\ldots$$

Nominal Process 4:

$$x_k = 0.7x_{k-1} - 0.3x_{k-2} + 0.2x_{k-3} + 0.1x_{k-4} + 0.05x_{k-5} + w_k \qquad k=\ldots,-1,0,1,\ldots$$

In all the four processes, $\{W_k\}$ is a zero-mean, unit variance, i.i.d. Gaussian process, such that W_k is independent of $\{X_j\}_{j=-\infty}^{k-1}$. We summarize the results for these processes, corresponding to the designed predictors in (29) and (30), for different values of m and for different values of ε, in the following tables and figures.

Tables 1, 2 and Fig. 1, Tables 3-5 and Figs. 2,3; Tables 6-8 and Figs. 4-6; Tables 9-11 and Figs. 7-11 correspond respectively to the nominal processes 1,2,3 and 4. Tables 1,3,6 and 9 give the asymptotic mean squared error (amse) at the nominal process 1,2,3 and 4 respectively, for the designed predictors in (29), and for different values of ε and m. Tables 1,4,7 and 10 give the amse at the nominal process 1,2,3 and 4 respectively, for the predictors in (30), and for different values of ε and m. Tables 2,5,8 and 11 give the breakdown points ε^* and $\bar{\varepsilon}^*$ of the predictors in (30), (and the predictor in (29) for m=1), for different values of ε and m. Figures 1 to 11 give the plots of the influence functions corresponding to the predictors in (30), for different values of ε and m and for all the four nominal processes.

From the above tables and figures, we make the following observations:

(1) When m=1, then for any ε, the amse at any of the nominal processes 1,2,3, or 4 is the same for the predictors in (29) and (30), as expected.

(2) If the nominal process is a pth order, auto-regressive, zero-mean, stationary, Gaussian, then for both the predictors in (29) and (30) and for all ε values and all $m \geq p$, the amse's at the nominal process coincide. Also, the breakdown points $\varepsilon^{*}_{G_2^{*m},m}$, $\bar{\varepsilon}^{*}_{G_2^{*m},m}$ and the influence functions coincide as well.

(3) For any $m \geq 1$, the amse at the nominal process μ_o viz. $e(\mu_o, G_i^{*m})$, i=1,2 converges to $e(\mu_o, m_o^*)$ as $\varepsilon \to 0$. As $\varepsilon \to 1$, $e(\mu_o, G_i^{*m})$ converges to σ^2. Except for the nominal process 1, $e(\mu_o, G_i^{*m})$ first increases with ε, exceeds σ^2 and then decreases, converging to σ^2 as $\varepsilon \to 1$. For the nominal process 1, $e(\mu_o, G_i^{*m})$ first increases with ε, exceeds σ^2, then decreases with ε, becomes less than σ^2 and again it increases, converging to σ^2 as $\varepsilon \to 1$.

(4) For most values of ε and m, the predictors $\{G_{2,n}^{*m}\}$ in (30) have smaller amse's at the nominal process, than the predictors $\{G_{1,n}^{*m}\}$. Also, for the predictors $\{G_{2,n}^{*m}\}$, the amse at the nominal process decreases with m; however, not necessarily in a monotone manner.

(5) The breakdown points $\varepsilon^{*}_{G_2^{*m},m}$ and $\bar{\varepsilon}_{G_2^{*m},m}$ are positive if and only if $e(\mu_o, G_2^{*m}) < \sigma^2$.

(6) The breakdown points $\varepsilon^{*}_{G_2^{*m},m}$ coincide for all values of ε and m and for all four nominal processes except in the case of the nominal process 3, with m=2 and the nominal process 4 with m=4. In those cases, and as observed by figures 5 and 10, the influence of the outliers is not maximum when $w = \infty$. Also, $\varepsilon_{G_2^{*m},m}{}^{*} \leq \bar{\varepsilon}^{*}_{G_2^{*m},m}$.

(7) Typically, for any ε, the breakdown point $\varepsilon^{*}_{G_2^{*m},m}$ is larger for large m than for small m.

(8) For any $m \geq 1$, starting from zero, the breakdown point $\varepsilon^{*}_{G_2^{*m},m}$ first increases with ε, and then it decreases with ε to zero. A plausible explanation is as follows: The breakdown point of the designed predictors $\{G_{2,n}^{*m}\}$ is determined mainly by two factors; $\sigma^2 - e(\mu_o, G_2^{*m})$, and λ_m.

For large breakdown point, the threshold λ_m should be small so that the influence of the outliers is small, and $\sigma^2 - e(\mu_o, G_2^{*m})$ should be large so that even in the presence of any level of contamination δ such that $\delta < \varepsilon^*_{G_2^{*m},m}$, the quantity $\sigma^2 - \sup_w e(\mu_{\delta,w,m}, G_2^{*m})$ remains positive. In our case, when ε is small, then both $\sigma^2 - e(\mu_o, G_2^{*m})$ and the threshold λ_m are large. When ε is large, then both λ_m and $\sigma^2 - e(\mu_o, G_2^{*m})$ are small.

(9) For any m≥1 and w≥0, the influence function $I_{G_2^{*m},m}[w]$ is a decreasing function of ε. Also, for any m≥1 and $\varepsilon>0$, the influence function is a bounded and continuous function of w. For $\varepsilon=0$, the influence function is not bounded.

(10) For the designed predictors in (30), the influence functions $I_{G_2^{*m},m}[w]$ are not always monotically increasing for positive w (see Figures 5,8,9 and 10). This is because the predictors $G_{2,n}^{*m}(w^m)$, when treated as functions of w only, are not monotonically increasing with w.

VI. APPENDIX

Proof of Theorem 1: Take any contaminating density $h_{n-1}(y_1^{n-1})$. Let $f_o(x_n, y_1^{n-1})$ be the density of (X_n, X_1^{n-1}). Let $f(x_n, y_1^{n-1})$ be the joint density of X_n and Y_1^{n-1}. Then,

$$f(x_n, y_1^{n-1}) = (1-\varepsilon) f_o(x_n, y_1^{n-1}) + \varepsilon f_o(x_n) h_{n-1}(y_1^{n-1}) \tag{A.1}$$

Also,

$$f(y_1^{n-1}) = (1-\varepsilon) f_o(y_1^{n-1}) + \varepsilon h_{n-1}(y_1^{n-1}) \tag{A.2}$$

Now, for any $g_n(y_1^{n-1})$ and any $h_{n-1}(y_1^{n-1})$, we have

$$E[\{X_n - g_n(Y_1^{n-1})\}^2] = \iint [x_n - g_n(y_1^{n-1})]^2 f(x_n, y_1^{n-1}) dx_n dy_1^{n-1}$$

$$= \int [\int [x_n - g_n(y_1^{n-1})]^2 f(x_n \mid y_1^{n-1}) dx_n] f(y_1^{n-1}) dy_1^{n-1} \tag{A.3}$$

By (A.3), for any given $f(x_n, y_1^{n-1})$, the optimum $g_n^*(y_1^{n-1})$ is: $g_n^*(y_1^{n-1}) = E[X_n \mid y_1^{n-1},]$ where the conditional density $f(x_n \mid y_1^{n-1})$ is used in evaluating $E[X_n \mid y_1^{n-1}]$. Also, maximizing (A.3) over $h_{n-1}(y_1^{n-1)}$ is equivalent to maximizing it over $f(y_1^{n-1})$ of (A.2). Hence, using $g_n^*(y_1^{n-1})$ in (A.3), by (A.1) and since $E[X_n] = 0$, we have for any $h_{n-1}(y_1^{n-1})$

$$\inf_{g_n(y_1^{n-1})} E[\{X_n - g_n(y_1^{n-1})\}^2]$$

$$= \int [\int \{x_n^2 - E^2[X_n \mid y_1^{n-1}]\} f(x_n \mid y_1^{n-1}) dx_n] f(y_1^{n-1}) dy_1^{n-1}$$

$$= \sigma^2 - \int [\int x_n f(x_n \mid y_1^{n-1}) dx_n]^2 f(y_1^{n-1}) dy_1^{n-1}] =$$

$$= \sigma^2 - \int \frac{[\int x_n f(x_n, y_1^{n-1}) dx_n]^2}{f(y_1^{n-1})} dy_1^{n-1}$$

$$= \sigma^2 - \int \frac{[(1-\varepsilon) f_o(y_1^{n-1}) \int x_n f_o(x_n \mid y_1^{n-1}) dx_n]^2}{f(y_1^{n-1})} dy_1^{n-1}$$

$$= \sigma^2 - \int \frac{[(1-\varepsilon) f_o(y_1^{n-1}) m_o(y^{n-1})]^2}{f(y_1^{n-1})} dy_1^{n-1} \tag{A.4}$$

Therefore, our objective now is

$$f^* = f^*(y_1^{n-1}) : \inf_{f(y_1^{n-1})} \int \frac{[(1-\varepsilon) f_o(y_1^{n-1}) m_o(y_1^{n-1})]^2}{f(y_1^{n-1})} dy_1^{n-1} \tag{A.5}$$

where $f(y_1^{n-1})$ is of the form (A.2), with constraints

a. $$f(y_1^{n-1}) - (1-\varepsilon) f_o(y_1^{n-1}) \geq 0 \quad \forall\, y^{n-1} \tag{A.6}$$

b. $$\int f(y_1^{n-1}) dy_1^{n-1} = 1 \tag{A.7}$$

We will use the Lagrange multiplier technique of the calculus of variations to determine f^*. The Lagrange functional without the constraint (A.6) is

$$J[f^*, \delta] = \int \frac{[(1-\varepsilon) f_o(y_1^{n-1}) m_o(y^{n-1})]^2}{f^*(y_1^{n-1}) + \delta p(y_1^{n-1})} dy_1^{n-1} - \alpha_n \int [f^*(y_1^{n-1}) + \delta p(y_1^{n-1})] dy_1^{n-1} \tag{A.8}$$

where α_n is the Lagrange multiplier. Hence

$$\left.\frac{\partial J}{\partial \delta}\right|_{\delta=0} = 0$$

$$\int dy_1^{n-1} p(y_1^{n-1}) \left\{ \left[\frac{(1-\varepsilon) f_o(y_1^{n-1}) m_o(y_1^{n-1})}{f^*(y_1^{n-1})} \right]^2 + \alpha_n \right\} = 0$$

$$\forall p : \int p(y_1^{n-1}) dy_1^{n-1} = 0 \tag{A.9}$$

By a fundamental theorem of the variational calculus,

$$\left| \frac{(1-\varepsilon) f_o(y_1^{n-1}) m_o(y_1^{n-1})}{f^*(y_1^{n-1})} \right| = \lambda_{n-1}$$

or

$$f^*(y_1^{n-1}) = (1-\varepsilon) f_o(y_1^{n-1}) \cdot \frac{|m_o(y_1^{n-1})|}{\lambda_{n-1}} \tag{A.10}$$

By (A.10) and the constraint (A.6), we have

$$f^*(y_1^{n-1}) = \begin{cases} (1-\varepsilon) f_o(y_1^{n-1}) \ , & f^*(y_1^{n-1}) = (1-\varepsilon) f_o(y_1^{n-1}) \\ (1-\varepsilon) f_o(y_1^{n-1}) \cdot \dfrac{|m_o(y_1^{n-1})|}{\lambda_{n-1}} \ , & f^*(y_1^{n-1}) > (1-\varepsilon) f_o(y_1^{n-1}) \end{cases} \tag{A.11}$$

Now, $f^*(y_1^{n-1}) > (1-\varepsilon) f_o(y_1^{n-1})$ iff

$$(1-\varepsilon) f_o(y_1^{n-1}) \frac{|m_o(y_1^{n-1})|}{\lambda_{n-1}} > (1-\varepsilon) f_o(y_1^{n-1})$$

or

$$|m_o(y_1^{n-1})| > \lambda_{n-1} \tag{A.12}$$

Hence, by (A.11) and (A.12)

$$f^*(y_1^{n-1}) = (1-\varepsilon)f_o(y_1^{n-1})\cdot\max\left\{1, \frac{|m_o(y_1^{n-1})|}{\lambda_{n-1}}\right\} \tag{A.13}$$

The positive constant λ_{n-1} is chosen so that $\int f^*(y_1^{n-1})dy_1^{n-1} = 1$.

The optimal predictor $g_n^*(y_1^{n-1})$ corresponds to f^*.

$$g_n^*(y_1^{n-1}) = E[X_n | y_1^{n-1}] = \int x_n f^*(x_n | y_1^{n-1})dx_n$$

$$= \frac{\int x_n f^*(x_n, y_1^{n-1})dx_n}{f^*(y_1^{n-1})}$$

$$= \frac{(1-\varepsilon)\int x_n f_o(x_n, y_1^{n-1})dx_n + \varepsilon\int x_n f(x_n)dx_n \cdot h_{n-1}^*(y_1^{n-1})}{f^*(y_1^{n-1})}$$

$$= \frac{(1-\varepsilon)\int x_n f_o(x_n, y_1^{n-1})dx_n}{-f^*(y_1^{n-1})} =$$

$$= \frac{(1-\varepsilon)f_o(y_1^{n-1})m_o(y_1^{n-1})}{f^*(y_1^{n-1})}$$

$$= \begin{cases} m_o(y_1^{n-1}) & \text{if } |m_o(y_1^{n-1})| \le \lambda_{n-1} \\ \lambda_{n-1} \cdot \dfrac{m_o(y_1^{n-1})}{|m_o(y_1^{n-1})|}, & \text{otherwise} \end{cases} \tag{A.14}$$

Proof of Theorem 2: The proof is based on the following lemma.

Lemma 1: Let N_1 and N_2 be two jointly normal random variables with mean zero and variances σ_1^2, σ_2^2 respectively. Let $E[N_1N_2] = \rho\sigma_1\sigma_2$. For any $\lambda>0$, let

$$h_\lambda(u) = u \cdot \min\{1, \frac{\lambda}{|u|}\} \tag{B.1}$$

Assume the integral of (65) exists for the tuple $(\lambda,\sigma_1,\sigma_2,\rho)$. Then,

$$E[h_\lambda(N_1)]^2 = \theta[\lambda,\sigma_1] \tag{B.2}$$

$$E[h_\lambda(N_1)h_\lambda(N_2)] = \beta[\lambda,\sigma_1,\sigma_2,\rho] \tag{B.3}$$

$$E[N_1 \cdot h_\lambda(N_2)] = \zeta[\lambda,\sigma_1,\sigma_2,\rho] \tag{B.4}$$

where θ, β, and ζ are defined in Section III.

Proof of the Lemma: (B.2) follows by direct computations. Define σ_c, γ_c as in (56). Then,

$$E[h_\lambda(N_1)h_\lambda(N_2)] = \int_{-\infty}^{\infty} h_\lambda(x) \cdot E[h_\lambda(N_2)\,|\,N_1 = x] \cdot \frac{\phi(\frac{x}{\sigma_1})}{\sigma_1}\, dx \tag{B.5}$$

Also,

$$E[h_\lambda(N_2)\mid N_1 = x] = \sigma_c\left[\phi\left(\frac{\gamma_c x+\lambda}{\sigma_c}\right) - \phi\left(\frac{\gamma_c x-\lambda}{\sigma_c}\right)\right]$$

$$+ (\gamma_c x+\lambda)\Phi\left(\frac{\gamma_c+\lambda}{\sigma_c}\right) - (\gamma_c x-\lambda)\Phi\left(\frac{\gamma_c x-\lambda}{\sigma_c}\right) - \lambda \qquad \text{(B.6)}$$

By Taylor's theorem,

$$\phi\left(\frac{\gamma_c x+\lambda'}{\sigma_c}\right) = \sum_{n=0}^{\infty} \frac{\phi^{(n)}\left(\frac{\lambda'}{\sigma_c}\right)}{n!}\left(\frac{\gamma_c}{\sigma_c}\right)^n \cdot x^n \qquad \lambda'=\lambda \text{ or } -\lambda \qquad \text{(B.7)}$$

$$(\gamma_c x + \lambda')\Phi\left(\frac{\gamma_c x+\lambda'}{\sigma_c}\right) = \lambda'\Phi\left(\frac{\lambda'}{\sigma_c}\right) + \gamma_c \cdot \Phi\left(\frac{\lambda'}{\sigma_c}\right)x$$

$$+ \lambda' \sum_{n=1}^{\infty} \phi^{(n-1)}\left(\frac{\lambda'}{\sigma_c}\right)\cdot\left(\frac{\gamma_c}{\sigma_c}\right)^n \cdot \frac{1}{n!}\cdot x^n$$

$$+ 2\sigma_c \sum_{n=2}^{\infty} \phi^{(n-2)}\left(\frac{\lambda'}{\sigma_c}\right)\left(\frac{\gamma_c}{\sigma_c}\right)^n \cdot \frac{1}{n!} x^n, \quad \lambda'=\lambda \text{ or } -\lambda \qquad \text{(B.8)}$$

Substituting (B.7), (B.8) in (B.6) and in turn (B.6) in (B.5), we get the desired results (B.3) using the Fubini's theorem, which is applicable if the integral of (65) exists. Similarly, assuming this integral exists, we get (B.4) using (B.6)-(B.8).

We will now give the proof of Theorem 2 for the mapping G_1^{*m}. The proof for G_2^{*m} follows in a similar manner. The asymptotic mean squared error $e(\mu_o, G_1^{*m})$ at the nominal source μ_o is

$$e(\mu_o, G_1^{*m}) = E[\{X_0 - G_1^{*m}(X_{-\infty}^{-1})\}^2] = \sigma^2 - 2E[X_0 \cdot G_1^{*m}(X_{-\infty}^{-1})]$$

$$+ E[\{G_1^{*m}(X_{-\infty}^{-1})\}^2] \tag{B.9}$$

By the definition of G_1^{*m} in (33),

$$e(\mu_o, G_1^{*m}) = \sigma^2 - 2 \sum_{i=-\infty}^{-1} \frac{d_i}{a_m^{(m)}} E\{X_o \cdot p_m[X_{i-m+1}^{i}]\}$$

$$+ \sum_{i=-\infty}^{-1} \sum_{j=-\infty}^{-1} \frac{d_i d_j}{[a_m^{(m)}]^2} E\{P_m[X_{i-m+1}^{i}] \cdot p_m[X_{j-m+1}^{j}]\} \tag{B.10}$$

The result in (63) follows from (B.10) and the definitions in (35) and (36).

Now, by the definitions of p_m in (31), h_λ in (B.1), and $Z_{1,i}$ in (39), $Z_{2,i}$ in (40), and the definition of g_{m+1}^{*} in (20), we get

$$R_{p_m}(0) = E\{p_m(X_{i-m+1}^{i})\}^2$$

$$= E\{g_{m+1}^{*}(X_{i-m+1}^{i}) - g_{m+1}^{*}(X_{i-m+1}^{i-1}, 0)\}^2$$

$$= E\{h_{\lambda_m}(Z_{1,i}) - h_{\lambda_m}(Z_{2,i})\}^2$$

$$= \sum_{t=1}^{2} E[h_{\lambda_m}(Z_{t,i})]^2 - 2E[h_{\lambda_m}(Z_{1,i}) h_{\lambda_m}(Z_{2i})] \tag{B.11}$$

$$R_{p_m}(j-i) = E\{p_m(X_{i-m+1}^{i}) p_m(X_{j-m+1}^{j})\}$$

$$= \sum_{l,t=1}^{2} (-1)^{l+t} E\{h_{\lambda_m}(Z_{l,i}) h_{\lambda_m}(Z_{t,j})\}, \quad i \neq j \tag{B.12}$$

$$R_{xp_m}(i) = E[X_0 \cdot p_m(X_{i-m+1}^{m})]$$

$$= E[X_0 h_{\lambda_m}(Z_{1,i})] - E[X_0 \cdot h_{\lambda_m}(Z_{2,i})], \quad i \leq -1 \tag{B.13}$$

The desired results, (61)-(67), follow directly from (B.11)-(B.13), Lemma 1, and (57).

Proof of Theorem 3:

(a) By the Minkowaski inequality and the definition of m_o^* in (30), we have:

$$\{e(\mu_o, G_1^{*m})\}^{1/2} = \{E[X_0 - G_1^{*m}(X_{-\infty}^{-1})]^2\}^{1/2}$$

$$= \{E[(X_0 - m_o^*(X_{-\infty}^{-1})) + (m_o^*(X_{-\infty}^{-1}) - G_1^{*m} \cdot (X_{-\infty}^{-1}))]^2\}^{1/2}$$

$$\leq \{E[X_0 - m_o^*(X_{-\infty}^{-1})]^2\}^{1/2} + \{E[m_o^*(X_{-\infty}^{-1}) - G_1^{*m}(X_{-\infty}^{-1})]^2\}^{1/2} =$$

$$= \sqrt{e(\mu_o, m_o^*)} + \left\{ E\left[\sum_{i=-\infty}^{-1} d_i \{X_i - \frac{p_m(X_{i-m+1}^{i})}{a_m^{(m)}}\} \right]^2 \right\}^{1/2} \tag{C.1}$$

By the definition of p_m (31) and J^m (78)

$$p_m(u_1^m) = u_m \cdot a_m^{(m)}, \quad \text{if } u_1^m \varepsilon J^m \tag{C.2}$$

and since $|p_m|$ is bounded from above by $2\lambda_m$, hence

$$|x_i - \frac{p_m(x_{i-m+1}^{i})}{a_m^{(m)}}| \leq \begin{cases} 0, & \text{if } x_{i-m+1}^{i} \varepsilon J^m \\ |x_i| + \dfrac{2\lambda_m}{|a_m^{(m)}|}, & \text{otherwise} \end{cases} \tag{C.3}$$

Therefore, by (C.1), (C.3) and the Minkowaski inequality, we obtain,

$$\left\{ E\left[\sum_{i=-\infty}^{-1} d_i \left\{ X_i - \frac{p_m(X^{i}_{i-m+1})}{a_m^{(m)}} \right\} \right]^2 \right\}^{1/2} \le$$

$$\le \sum_{i=-\infty}^{-1} |d_i| \left\{ E\left[X_i - \frac{p_m(X^{i}_{i-m+1})}{a_m^{(m)}} \right]^2 \right\}^{1/2}$$

$$\le \sum_{i=-\infty}^{-1} |d_i| \{E([|X_i| + \frac{2\lambda_m}{|a_m^{(m)}|}]^2 \cdot (1-1_{J^m}(X^{i}_{i-m+1})))\}^{1/2} \tag{C.4}$$

By the stationarity of the process μ_o,{·} of (C.4) equals H_m. Hence, by (C.1), (C.4) and (79), we conclude:

$$c(\mu_o, G_1^{*m}) \le [(c(\mu_o, m_o^{*}))^{1/2} + D^{*} \cdot \sqrt{H_m}]^2$$

(b) By (a), it suffices to prove that

$$\lim_{\varepsilon \to 0} H_m = 0 \tag{C.5}$$

By the definition of J^m in (78), we have,

$$H_m = E[(|X_m| + \frac{2\lambda_m}{|a_m^{(m)}|})^2 (1-1_{J^m}(X_1^m))]$$

$$\leq E[(|X_m| + \frac{2\lambda_m}{|a_m^{(m)}|})^2(1-1_{J_1^m}(X_1^m))] + E[(|X_m| + \frac{2\lambda_m}{|a_m^{(m)}|})^2(1-1_{J_2^m}(X_1^m))] \tag{C.6}$$

where

$$J_1^m = \{y_1^m : |\sum_{j=1}^{m} a_j^{(m)} y_j| < \lambda_m\} \tag{C.7a}$$

$$J_2^m = \{y_1^m : |\sum_{j=1}^{m-1} a_j^{(m)} y_j| < \lambda_m\} \tag{C.7b}$$

Now, as $\varepsilon \to 0$, $\lambda_m \to \infty$ and hence, $1-1_{J_1^m}(x_1^m)$ decreases to zero as $\varepsilon \to 0$ for all x_1^m. Since X_m has finite variance, hence by the dominated convergence theorem we have:

$$\lim_{\varepsilon \to 0} E[X_m^2(1-1_{J_1^m}(X_1^m))] = 0 \tag{C.8}$$

Also, by the definition of J_1^m in (C.7a), we conclude,

$$\lambda_m^2 E[1-1_{J_1^m}(X_1^m)] = E[\lambda_m^2 \cdot (1-1_{J_1^m}(X_1^m))]$$

$$\leq E[|\sum_{j=1}^{m} a_j^{(m)} X_j|^2 \cdot (1-1_{J_1^m}(X_1^m))]$$

$$\leq 2^{m-1} \sum_{j=1}^{m} |a_j^{(m)}|^2 E[X_j^2 \cdot (1-1_{J_1^m}(X_1^m))] \tag{C.9}$$

Therefore, as in (C.8), we have,

$$\lim_{\varepsilon \to 0} \frac{4}{|a_m^{(m)}|^2} \cdot \lambda_m^2 \cdot E[1-1_{J_1^m}(X_1^m)] = 0 \tag{C.10}$$

Similarily,

$$\lim_{\varepsilon\to 0} \frac{4}{|a_m^{(m)}|} \cdot \lambda_m E[\,|X_m|(1-1_{J_1^m}(X_1^m))]$$

$$\le \lim_{\varepsilon\to 0} \frac{4}{|a_m^{(m)}|} \cdot \sum_{j=1}^{m} |a_j^{(m)}| \cdot E[\,|X_m X_j|(1-1_{J_1^m}(X_1^m))] = 0 \tag{C.11}$$

Hence, by (C.8), (C.10), and (C.11), the first term on the right of (C.6) goes to zero as $\varepsilon\to 0$. By a similar analysis, the second term on the right of (C.6) goes to zero as $\varepsilon\to 0$ and therefore, (C.5) holds.

REFERENCES

[1] Boente, G., Fraiman, R. and Yohai, V. (1982), "Qualitative robustness for general stochastic processes", Technical Report #26, Oct., Dept. of Statistics, Univ. of Washington, Seattle, WA.

[2] Cox, D. (1981), "Metrics on stochastic processes and qualitative robustness", TR#3, Dept. of Statistics, Univ. of Washington, Seattle, WA.

[3] Franke, J. (1985), "Minimax robust prediction of discrete time series", Z. Wahrsheinlichkeitstheorie Verw. Geb., 68, pp. 337-364.

[4] Franke, J. and Poor, H.V. (1984), "Minimax robust filtering andfinite length robust predictors", Robust and Nonlinear Time Series Analysis, in Lecture Notes in Statistics, 26, Springer-Verlag.

[5] Ferguson, T.S. (1967), "Mathematical Statistics", Academic Press, New York.

[6] Grenander, U. and Szego, G. (1958), "Toeplitz Forms and Their Applications", Univ. of California Press, Berkeley.

[7] Hampel, F.R. (1971), "A general qualitative definition of robustness", Annals of Math. Stat., 42, 1887-1896.

[8] Hosoya, Y. (1978), "Robust linear extrapolations of second order stationary processes", Annals of Prob., Vol. 6, #4, 574-584.

[9] Huber, P.J. (1981), "Robust Statistics", Wiley, New York.

[10] Kassam, S.A. and Poor, H.V. (1985), "Robust techniques for signal processing: A survey", IEEE Proceedings, vol. 73, no. 3, 433-479.

[11] Kolmogorov, A. (1941), "Interpolation and extrapolation", Bull. Acad. Sci., USSR, Ser. Math., Vol. 5, 3-14.

[12] Kolmogorov, A. (1941), "Stationary sequences in Hilbert Space", Bull. Math. Univ. Moscow, 2-40.

[13] Papantoni-Kazakos, P. and Gray, R.M. (1979), "Robustness of estimators on stationary observations", Annals of Probability, 7, 9089-1002.

[14] Papantoni-Kazakos, P. (1987), "Qualitative robustness in time-series", Information and Computation, 72, 239-269.

[15] Papantoni-Kazakos, P. (1984), "A game theoretic approach to robust filtering", Info. and Control, 60, 168-191.

[16] Papantoni-Kazakos, P. (1984), "Some aspects of qualitative robustness in time series", Robust and Nonlinear Time Series Analysis, in Lecture Notes in Statistics, 26, 218-230, Springer-Verlag.

[17] Tsaknakis, H., Kazakos, D. and Papantoni-Kazakos, P. (1986), "Robust prediction and interpolation for vector stationary processes", Probability Theory and Related Fields, 72, 589-602.

[18] Tsaknakis, H. (1986), "Robust Methods in Time Seeries", PhD Dissertation, Univ. of Connecticut, Storrs.

[19] Wiener, N. (1949), "Extrapolation, Interpolation and Smoothing of Stationary Time Series", MIT Press, Cambridge, Mass.

[20] Waiters, P., "An Introduction to Ergodic Theory", Graduate Text in Math., #79, Springer-Verlag.

[21] Ash, R., Probability Theory, Wiley.

ε \ m	m=1	m=2	m=3	m=4	m=5	m=6
0.0001	1.011	1.011	1.011	1.011	1.011	1.011
0.001	1.049	1.049	1.049	1.049	1.049	1.049
0.005	1.115	1.115	1.115	1.115	1.115	1.115
0.01	1.161	1.161	1.161	1.161	1.161	1.161
0.025	1.257	1.257	1.257	1.257	1.257	1.257
0.05	1.357	1.357	1.357	1.357	1.357	1.357
0.075	1.435	1.435	1.435	1.435	1.435	1.435
0.10	1.523	1.523	1.523	1.523	1.523	1.523
0.125	1.627	1.627	1.627	1.627	1.627	1.627
0.15	1.739	1.739	1.739	1.739	1.739	1.739
0.20	1.957	1.957	1.957	1.957	1.957	1.957
0.30	2.221	2.221	2.221	2.221	2.221	2.221
0.50	2.069	2.069	2.069	2.069	2.069	2.069
0.70	1.782	1.782	1.782	1.782	1.782	1.782
0.90	1.834	1.834	1.834	1.834	1.834	1.834
0.99	1.947	1.947	1.947	1.947	1.947	1.947

Table 1: Asymptotic mse, $e(\mu_o, G_1^{*m})$ in (63) of the predictors in (29) at the nominal process 1. For this process $e(\mu_o, G_1^{*m}) = e(\mu_o, G_2^{*m}) \forall \varepsilon, \forall m$. Also, $e(\mu_o, m_o^*) = 1.00$ and the nominal process variance, $\sigma^2 = 1.9608$.

ε \ m	m=1	m=2	m=3	m=4	m=5	m=6
0.0001	.0863	.0863	.0863	.0863	.0863	.0863
0.001	.1204	.1204	.1204	.1204	.1204	.1204
0.005	.1584	.1584	.1584	.1584	.1584	.1584
0.01	.1803	.1803	.1803	.1803	.1803	.1803
0.025	.2137	.2137	.2137	.2137	.2137	.2137
0.05	.2432	.2432	.2432	.2432	.2432	.2432
0.075	.2595	.2595	.2595	.2595	.2595	.2595
0.10	.2594	.2594	.2594	.2594	.2594	.2594
0.125	.2387	.2387	.2387	.2387	.2387	.2387
0.15	.1937	.1937	.1937	.1937	.1937	.1937
0.20	.0055	.0055	.0055	.0055	.0055	.0055
0.30	0	0	0	0	0	0
0.50	0	0	0	0	0	0
0.70	.7535	.7535	.7535	.7535	.7535	.7535
0.90	.9536	.9536	.9536	.9536	.9536	.9536
0.99	.9956	.9956	.9956	.9956	.9956	.9956

Table 2: Breakdown point $\varepsilon^*_{G_2^{*m},m}$ in (15) of the predictors in (30) and for the nominal process 1. For this process, $\varepsilon^*_{G_2^{*m},m} = \bar{\varepsilon}^*_{G_2^{*m},m}$ $\forall\varepsilon, \forall m$.

$\varepsilon \backslash m$	m=1	m=2	m=3	m=4	m=5	m=6
0.0001	1.011	1.209	1.209	1.209	1.209	1.209
0.001	1.051	1.236	1.236	1.236	1.236	1.236
0.005	1.133	1.288	1.288	1.288	1.288	1.288
0.01	1.196	1.327	1.327	1.327	1.327	1.327
0.025	1.326	1.407	1.407	1.407	1.407	1.407
0.05	1.471	1.503	1.503	1.503	1.503	1.503
0.075	1.575	1.578	1.578	1.578	1.578	1.578
0.10	1.654	1.639	1.639	1.639	1.639	1.639
0.125	1.715	1.690	1.690	1.690	1.690	1.690
0.15	1.763	1.733	1.733	1.733	1.733	1.733
0.20	1.828	1.797	1.797	1.797	1.797	1.797
0.30	1.833	1.866	1.866	1.866	1.866	1.866
0.50	1.842	1.859	1.859	1.859	1.859	1.859
0.70	1.727	1.748	1.748	1.748	1.748	1.748
0.90	1.602	1.608	1.608	1.608	1.608	1.608
0.99	1.553	1.553	1.553	1.553	1.553	1.553

Table 3 Asymptotic mse $e(\mu_0, G_1^{*m})$ in (63) of the predictors in (29) at the nominal process 2. Hence, $e(\mu_0, m_0^*) = 1.00$ and the nominal process variance, $\sigma^2 = 1.5476$.

ε \ m	m=1	m=2	m=3	m=4	m=5	m=6
0.0001	1.011	1.006	1.006	1.006	1.006	1.006
0.001	1.051	1.029	1.029	1.029	1.029	1.029
0.005	1.133	1.074	1.074	1.074	1.074	1.074
0.01	1.196	1.109	1.109	1.109	1.109	1.109
0.025	1.326	1.185	1.185	1.185	1.185	1.185
0.05	1.471	1.278	1.278	1.278	1.278	1.278
0.075	1.575	1.353	1.353	1.353	1.353	1.353
0.10	1.654	1.414	1.414	1.414	1.414	1.414
0.125	1.715	1.465	1.465	1.465	1.465	1.465
0.15	1.763	1.508	1.508	1.508	1.508	1.508
0.20	1.828	1.575	1.575	1.575	1.575	1.575
0.30	1.883	1.653	1.653	1.653	1.653	1.653
0.50	1.842	1.685	1.685	1.685	1.685	1.685
0.70	1.727	1.642	1.642	1.642	1.642	1.642
0.90	1.602	1.577	1.577	1.577	1.577	1.577
0.99	1.553	1.550	1.550	1.550	1.550	1.550

Table 4: Asymptotic mse, $e(\mu_0, G_2^{*m})$ in (64) of the predictors in (30) at the monial process 2. Here, $e(\mu_0, m_0^*) = 1.00$ and the nominal process variance $\sigma^2 = 1.5476$.

ε \ m	m=1	m=2	m=3	m=4	m=5	m=6
0.0001	.0542	.0862	.0862	.0862	.0862	.0862
0.001	.0741	.1202	.1202	.1202	.1202	.1202
0.005	.0894	.1561	.1561	.1561	.1561	.1561
0.01	.0922	.1747	.1747	.1747	.1747	.1747
0.025	.0803	.1974	.1974	.1974	.1974	.1974
0.05	.0380	.2009	.2009	.2009	.2009	.2009
0.075	0	.1856	.1856	.1856	.1856	.1856
0.10	0	.1578	.1578	.1578	.1578	1578
0.125	0	.1192	.1192	.1192	.1192	.1192
0.15	0	.0693	.0693	.0693	.0693	.0693
0.20	0	0	0	0	0	0
0.30	0	0	0	0	0	0
0.50	0	0	0	0	0	0
0.70	0	0	0	0	0	0
0.90	0	0	0	0	0	0
0.99	0	0	0	0	0	0

Table 5: Breakdown point $\varepsilon^{*}_{G_2^{*m},m}$ in (15) of the predictors in (30) and for the nominal process 2. For this process, $\varepsilon^{*}_{G_2^{*},m} = \varepsilon^{*}_{G_2^{*m},m}$ $\forall\varepsilon$, $\forall m$.

ε \ m	m=1	m=2	m=3	m=4	m=5	m=6
0.0001	1.012	1.083	1.233	1.233	1.233	1.233
0.001	1.054	1.117	1.264	1.264	1.264	1.264
0.005	1.139	1.185	1.325	1.325	1.325	1.325
0.01	1.203	1.237	1.370	1.370	1.370	1.370
0.025	1.333	1.345	1.461	1.461	1.461	1.461
0.05	1.482	1.469	1.568	1.568	1.568	1.568
0.075	1.593	1.561	1.649	1.649	1.649	1.649
0.10	1.678	1.633	1.716	1.716	1.716	1.716
0.125	1.745	1.691	1.771	1.771	1.771	1.771
0.15	1.799	1.737	1.816	1.816	1.816	1.816
0.20	1.874	1.805	1.885	1.885	1.885	1.885
0.30	1.942	1.874	1.957	1.957	1.957	1.957
0.50	1.908	1.875	1.938	1.938	1.938	1.938
0.70	1.786	1.786	1.811	1.811	1.811	1.811
0.90	1.650	1.653	1.657	1.657	1.657	1.657
0.99	1.596	1.596	1.597	1.597	1.597	1.597

Table 6: Asymptotic mse, $e(\mu_0, G_1^{*m})$ in (63) of the predictors in (29) at the nominal process 3. Here $e(\mu_0, m_0^*) = 1.00$ and the nominal process variance, $\sigma^2 = 1.5909$.

ε \ m	m=1	m=2	m=3	m=4	m=5	m=6
0.0001	1.012	1.057	1.007	1.007	1.007	1.007
0.001	1.054	1.084	1.031	1.031	1.031	1.031
0.005	1.139	1.139	1.078	1.078	1.078	1.078
0.01	1.203	1.181	1.115	1.115	1.115	1.115
0.025	1.333	1.271	1.193	1.193	1.193	1.193
0.05	1.482	1.378	1.291	1.291	1.291	1.291
0.075	1.593	1.461	1.370	1.370	1.370	1.370
0.10	1.678	1.528	1.435	1.435	1.435	1.435
0.125	1.745	1.583	1.491	1.491	1.491	1.491
0.15	1.799	1.629	1.538	1.538	1.538	1.538
0.20	1.874	1.698	1.611	1.611	1.611	1.611
0.30	1.942	1.776	1.698	1.698	1.698	1.698
0.50	1.908	1.799	1.736	1.736	1.736	1.736
0.70	1.786	.1734	1.691	1.691	1.691	1.691
0.90	1.650	1.636	1.622	1.622	1.622	1.622
0.99	1.595	1.596	1.594	1.594	1.594	1.594

Table 7: Asymptotic mse, $e(\mu_0, G_2^{*m})$ in (64) of the predictors in (30) at the nominal process 3. Here, $e(\mu_0, m_0^*) = 1.00$ and $\sigma^2 = 1.5909$.

ε \ m	m=1	m=2	m=3	m=4	m=5	m=6
0.0001	.0524	.0975 .1556	.0862	.0862	.0862	.0862
0.001	.0712	.1314 .2036	.1203	.1203	.1203	.1203
0.005	0.862	.1648 .2474	.1565	.1565	1565	.1565
0.01	.0900	.1800 .2667	.1756	.1756	.1756	.1756
0.025	.0824	.1947 .2832	.1999	.1999	.1999	.1999
0.05	.0479	.1795 .2677	.2061	.2061	.2061	.2061
0.075	0	.1514 .2236	.1934	.1934	.1934	.1934
0.10	0	.0955 .1500	.1684	.1684	.1684	.1684
0.125	0	.0170 .0280	.1326	.1326	.1326	.1326
0.15	0	0	.0858	.0858	.0858	.0858
0.20	0	0	0	0	0	0
0.30	0	0	0	0	0	0
0.50	0	0	0	0	0	0
0.70	0	0	0	0	0	0
0.90	0	0	0	0	0	0
0.99	0	0	0	0	0	0

Table 8: Breakdown point $\varepsilon^*_{G_2^{*m},m}$ in (15) of the predictors in (30) and for the nominal process 3. For this process, $\varepsilon^*_{G_2^{*m},m} = \varepsilon^*_{G_2^{*m},m}$ and the lower value is $\varepsilon^*_{G_2^{*m},m}$.

ε \ m	m=1	m=2	m=3	m=4	m=5	m=6
0.0001	1.014	1.072	1.265	1.206	1.194	1.194
0.001	1.062	1.112	1.304	1.250	1.240	1.240
0.005	1.155	1.189	1.375	1.331	1.323	1.323
0.01	1.223	1.248	1.426	1.388	1.381	1.381
0.025	1.363	1.372	1.528	1.501	1.496	1.496
0.05	1.532	1.522	1.648	1.630	1.627	1.627
0.075	1.665	1.640	1.747	1.735	1.734	1.734
0.10	1.775	1.737	1.835	1.833	1.833	1.833
0.125	1.866	1.819	1.917	1.924	1.927	1.927
0.15	1.943	1.887	1.991	2.010	2.015	2.015
0.20	2.058	1.992	2.114	2.155	2.164	2.164
0.30	2.182	2.110	2.256	2.324	2.338	2.338
0.50	2.192	2.143	2.241	2.291	2.302	2.302
0.70	2.070	2.057	2.070	2.078	2.080	2.080
.09	1.920	1.921	1.909	1.904	1.903	1.903
0.99	1.860	1.860	1.859	1.858	1.858	1.858

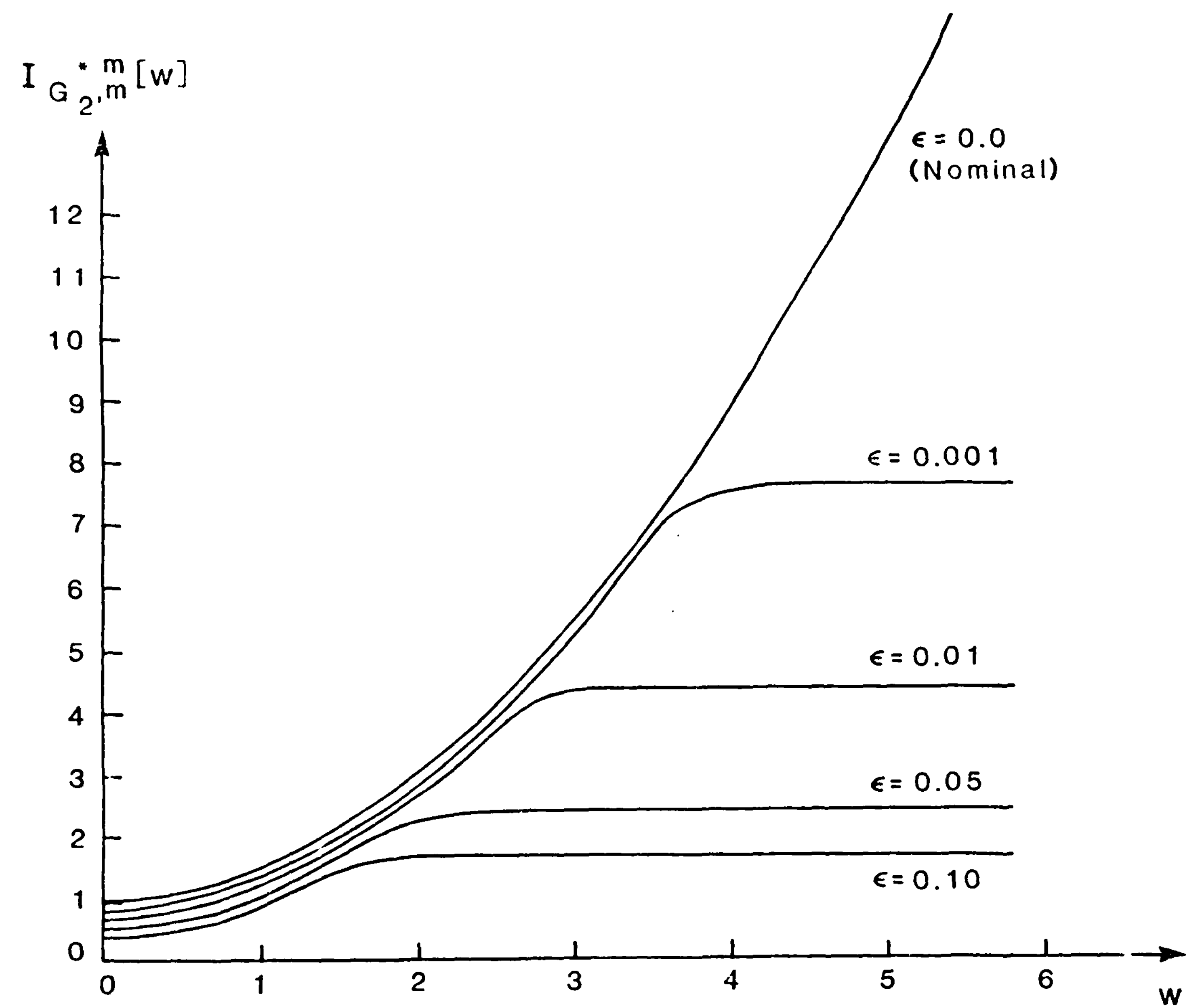
Table 9: Asymptotic msc. $e(\mu_o, G_1^{*m})$ in (63) of the predictors in (29) at the nominal process 4. Here, $e(\mu_o, m_o^*) = 1.00$ and the nominal process variance, $\sigma^2 = 1.8543$.

ε \ m	m=1	m=2	m=3	m=4	m=5	m=6
0.0001	1.014	1.130	1.020	1.001	1.010	1.010
0.001	1.062	1.162	1.050	1.034	1.044	1.044
0.005	1.155	.1222	1.105	1.095	1.105	1.105
0.01	1.223	1.269	1.146	1.138	1.148	1.148
0.025	1.363	1.367	1.234	1.229	1.238	1.238
0.05	1.532	1.491	1.344	1.341	1.346	1.346
0.075	1.665	1.593	1.439	1.436	1.438	1.438
0.10	1.775	1.679	1.528	1.527	1.526	1.526
0.125	1.866	1.754	1.610	1.615	1.612	1.612
0.15	1.943	1.818	1.685	1.697	1.692	1.692
0.20	2.058	1.920	1.813	1.840	1.832	1.832
0.30	2.182	2.044	1.972	2.016	2.000	2.000
0.50	2.192	2.113	2.017	X	X	X
0.70	2.070	2.065	1.935	X	X	X
0.90	1.920	1.928	.1869	X	X	X
0.99	1.860	1.861	1.855	X	X	X

Table 10: Asymptotic mse, $e(\mu_o, G_2^{*m})$ in (64) of the predictors in (30) at the nominal process 4. Here, $e(\mu_o, m_o^*) = 1.00$ and $\sigma^2 = 1.8543$. The symbol X denotes that $e(\mu_o, G_2^{*m})$ could not be computed because the condition in (65) is not satisfied.

ε \ m	m=1	m=2	m=3	m=4	m=5	m=6
0.0001	.0630	.0497	.0862	.0949 .1511	.0863	.0863
0.001	.0867	.0707	.1184	.1315 .2045	.1204	.1204
0.005	.1094	.0939	.1532	.1725 .2615	.1580	.1580
0.01	.1196	.1066	.1726	.1955 .2924	.1793	.1793
0.025	.1274	.1228	.2012	.2276 .3366	.2112	.2112
0.05	.1148	.1257	.2200	.2499 .3652	.2334	.2334
0.075	.0859	.1144	.2231	.2554 .3696	.2379	.2379
0.10	.0446	.0933	.2149	.2424 .3551	.2280	.2280
0.125	0	.0641	.1962	.2173 .3210	.2039	.2039
0.15	0	.0275	.1665	.1755 .2635	.1648	.1648
0.20	0	0	.0647	.0253 .0414	.0346	.0346
0.30	0	0	0	0	0	0
0.50	0	0	0	X	X	X
0.70	0	0	0	X	X	X
.090	0	0	0	X	X	X
0.99	0	0	0	X	X	X

Table II: Breakdown point $\varepsilon^*_{G_2^{*m},m}$ in (15) of the predictor in (30) and for the nominal process 4. For this process, $\underline{\varepsilon}^*_{G_2^{*m},m} = \bar{\varepsilon}^*_{G_2^{*m},m}$ $\forall\varepsilon$, $\forall m$, $m \neq 4$. For m=4, the upper value is $\underline{\varepsilon}^*_{G_2^{*m},m}$ and the lower value is $\bar{\varepsilon}^*_{G_2^{*m},m}$.



<u>Figure 1</u>. Influence functions $I^{*m}_{G_2,m}[w]$ in (96) for the nominal process 1 and m=1. For this process, $I^{*m}_{G_2,m}[w]$ are identical for all $m \geq 1$.

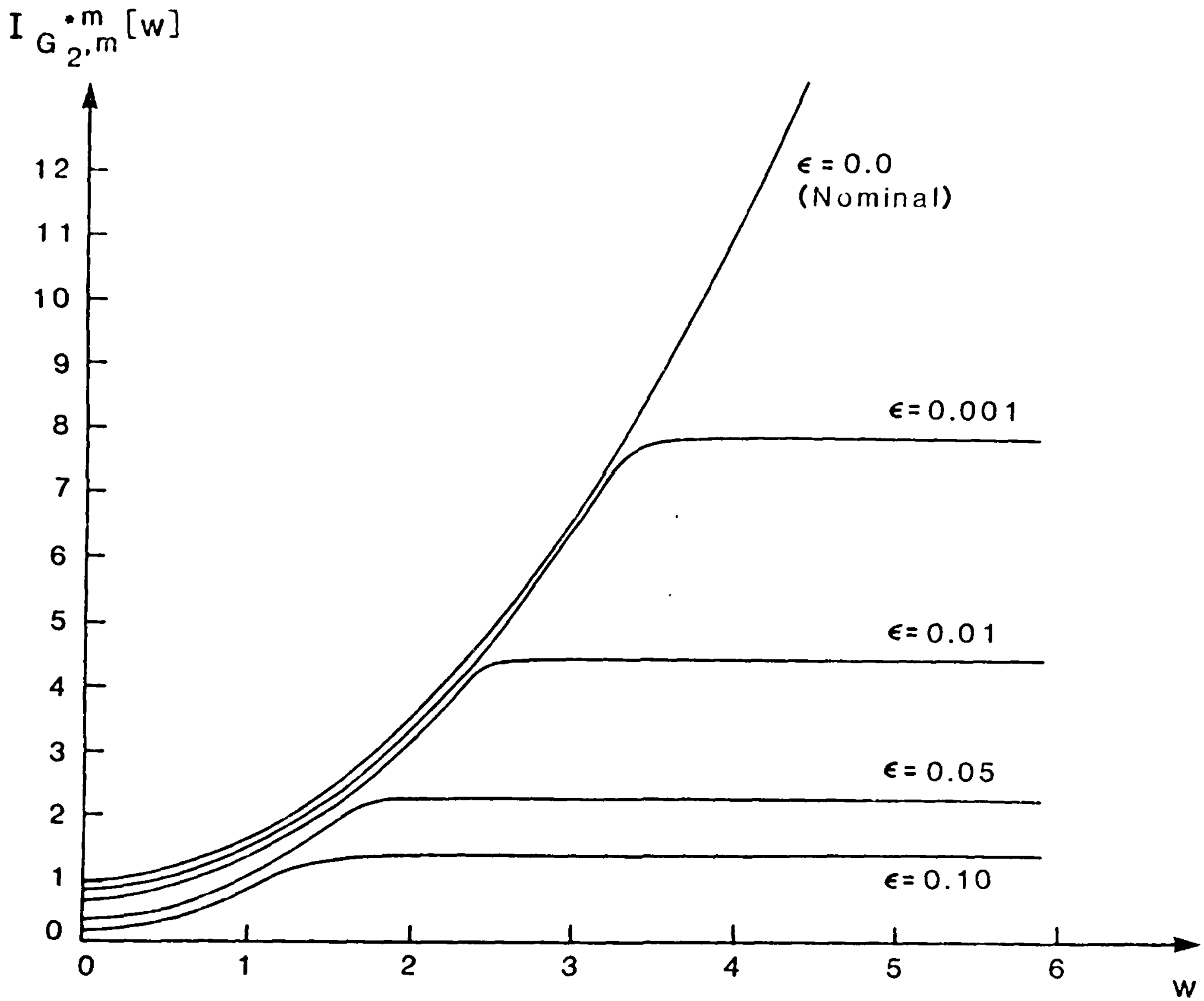


Figure [illegible]. Influence functions $I_{G_2^{*m},m}[w]$ in (96) for the nominal process [illegible] and m=1.

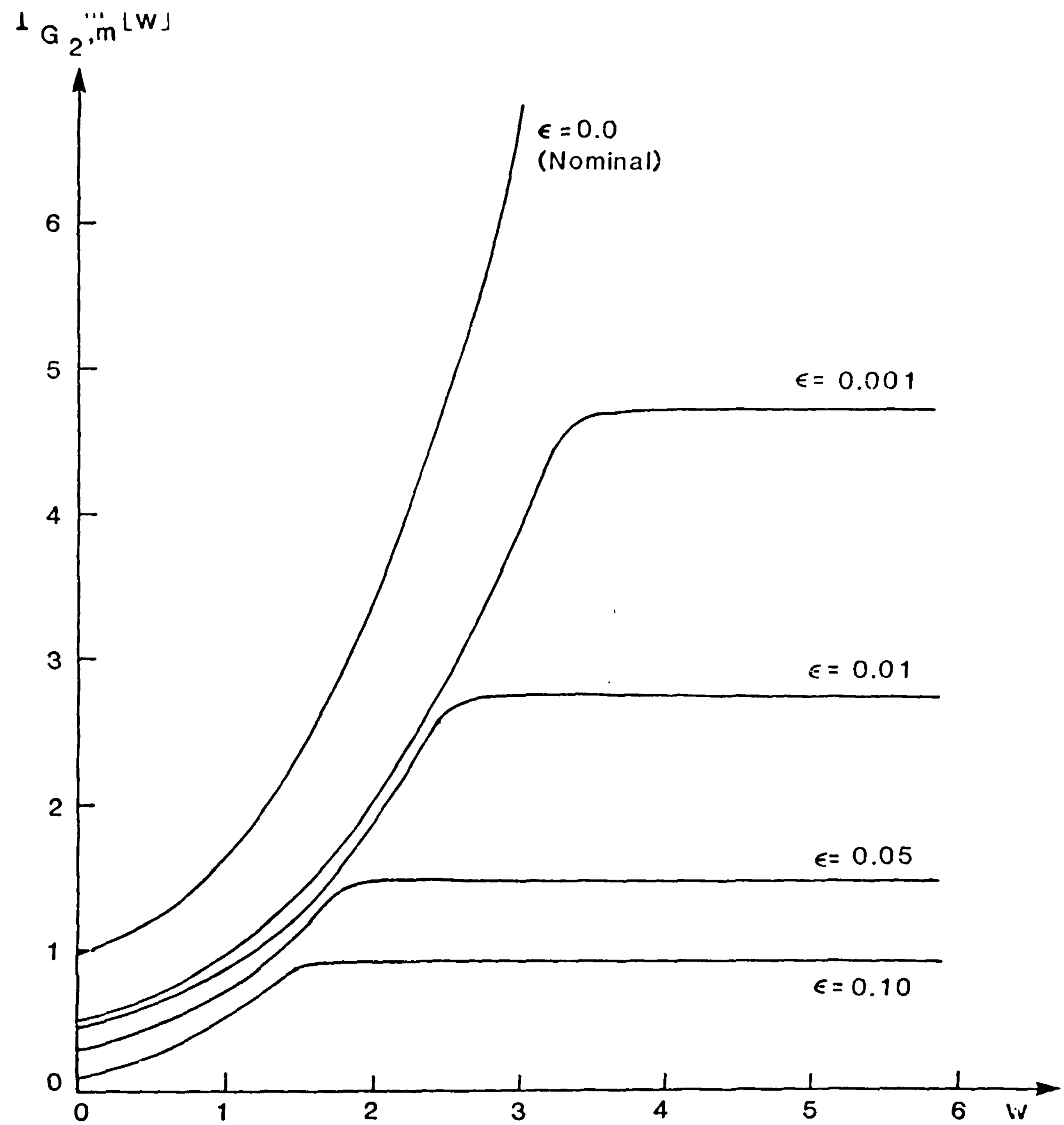


Figure 3. Influence functions $I_{G_2^*m,m}[w]$ in (96) for the nominal process 2 and m=2. For this process $I_{G^*m,m}[w]$ are identical for m≥

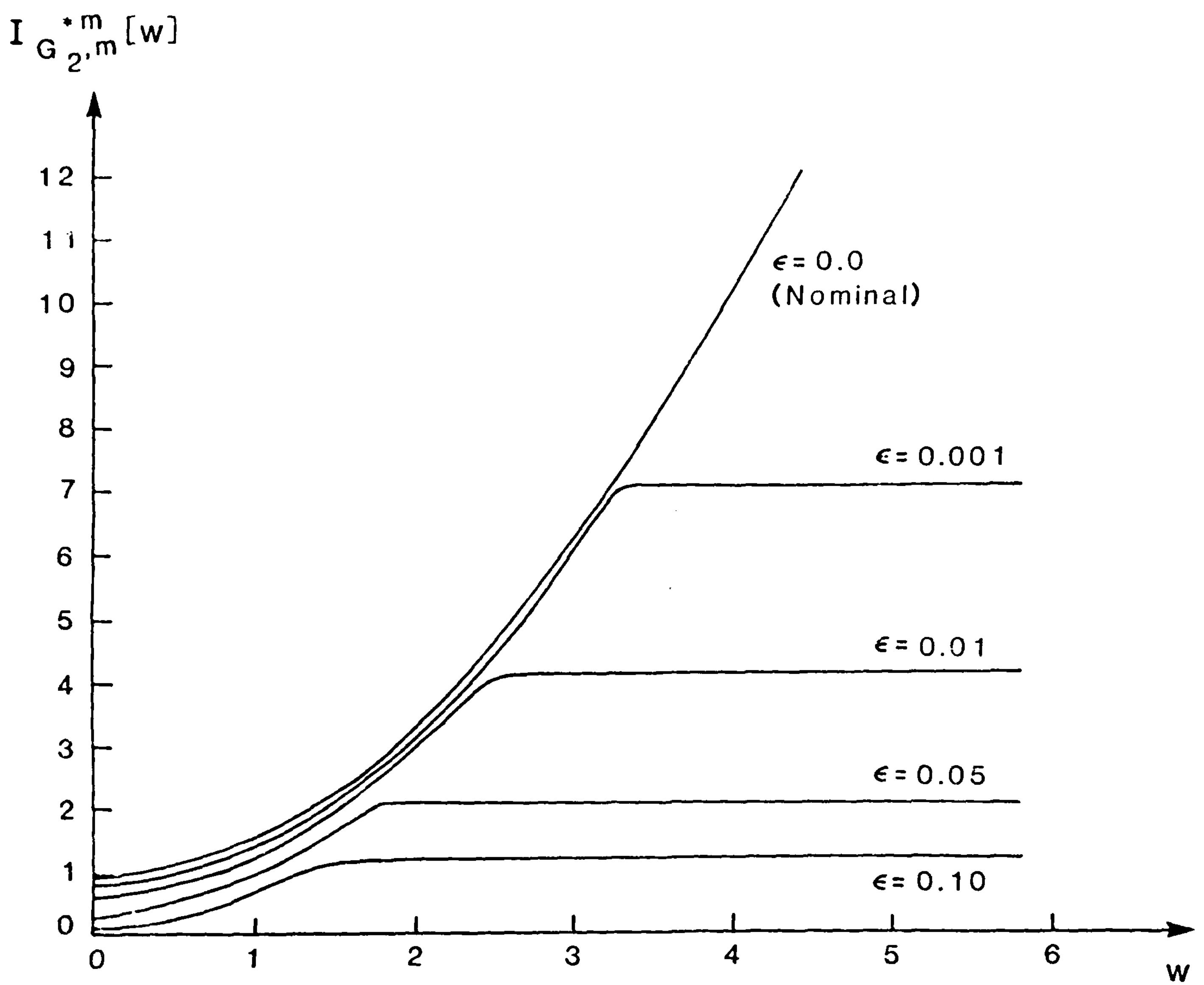


Figure 4. Influence functions $I_{G_2,m}^{*m}[w]$ in (96) for the nominal process 3 and m=1.

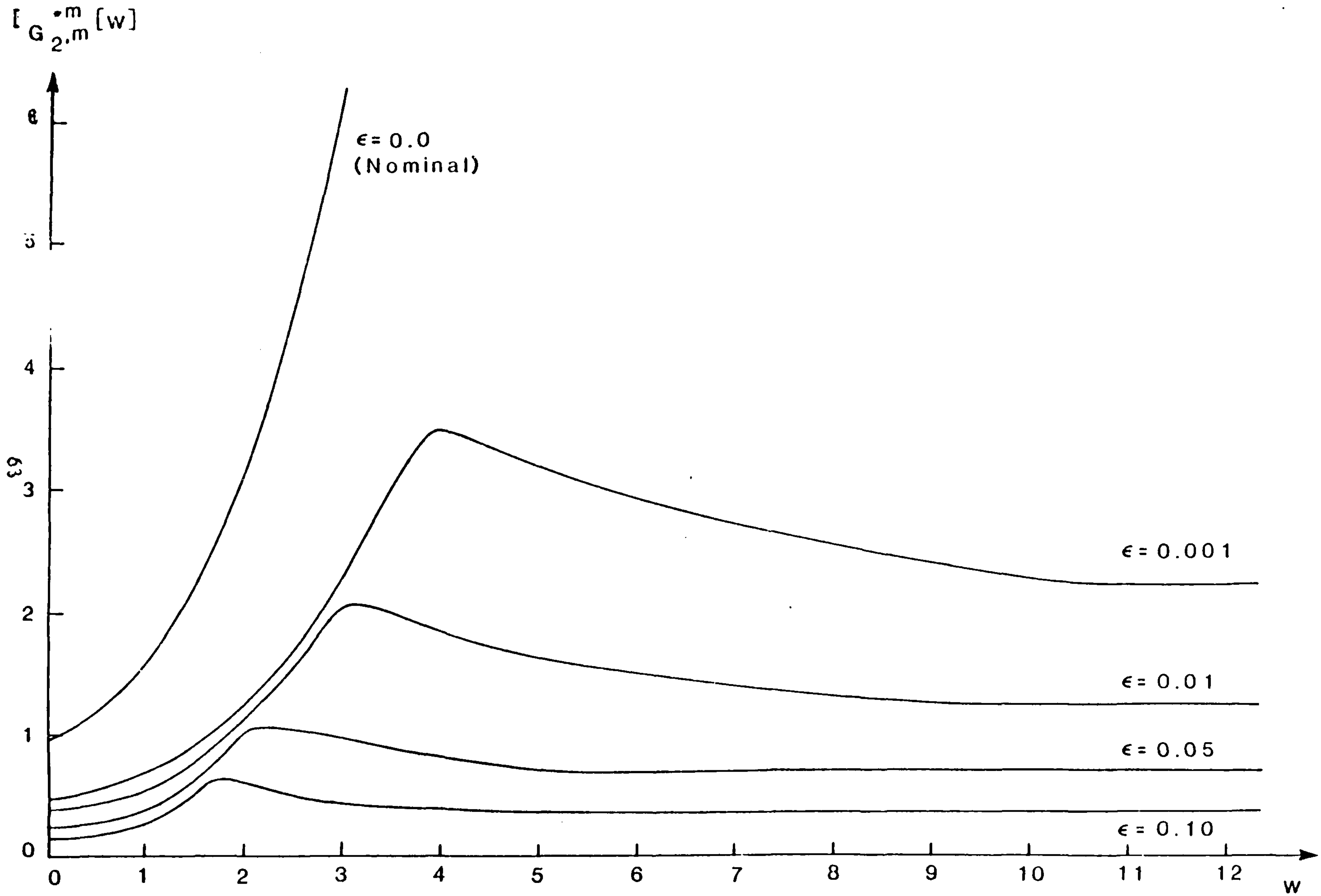


Figure 5. Influence functions $I_{G_2^{*m},m}[w]$ in (96) of the nominal process 3 and m=2.

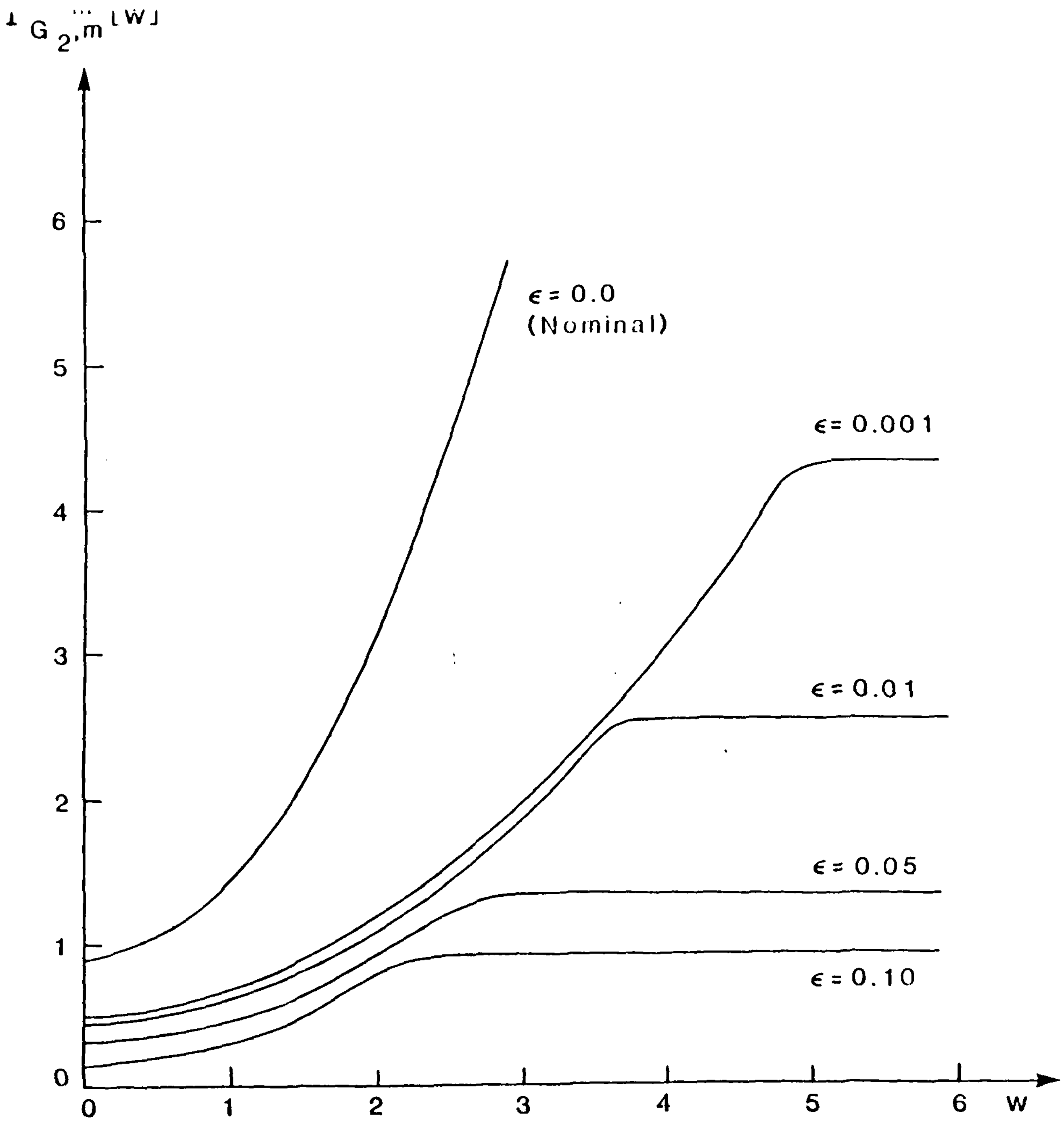


Figure 6. Influence functions $I_{G^*_{2,m}}[w]$ in (...) of the nominal ...

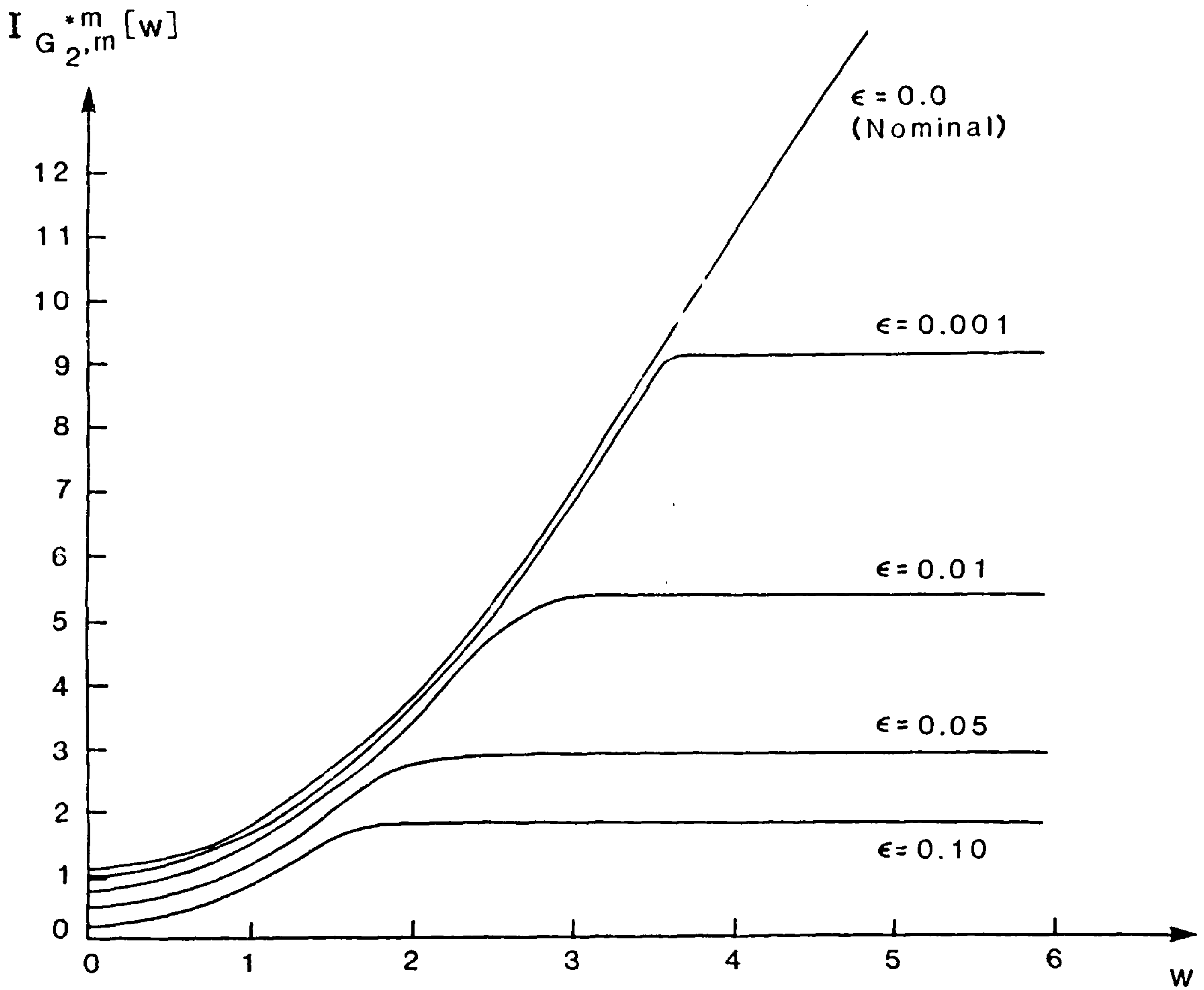


Figure 7. Influence functions $I_{G_2^{*m},m}[w]$ in [illegible] of the nominal process [illegible] and m=1.

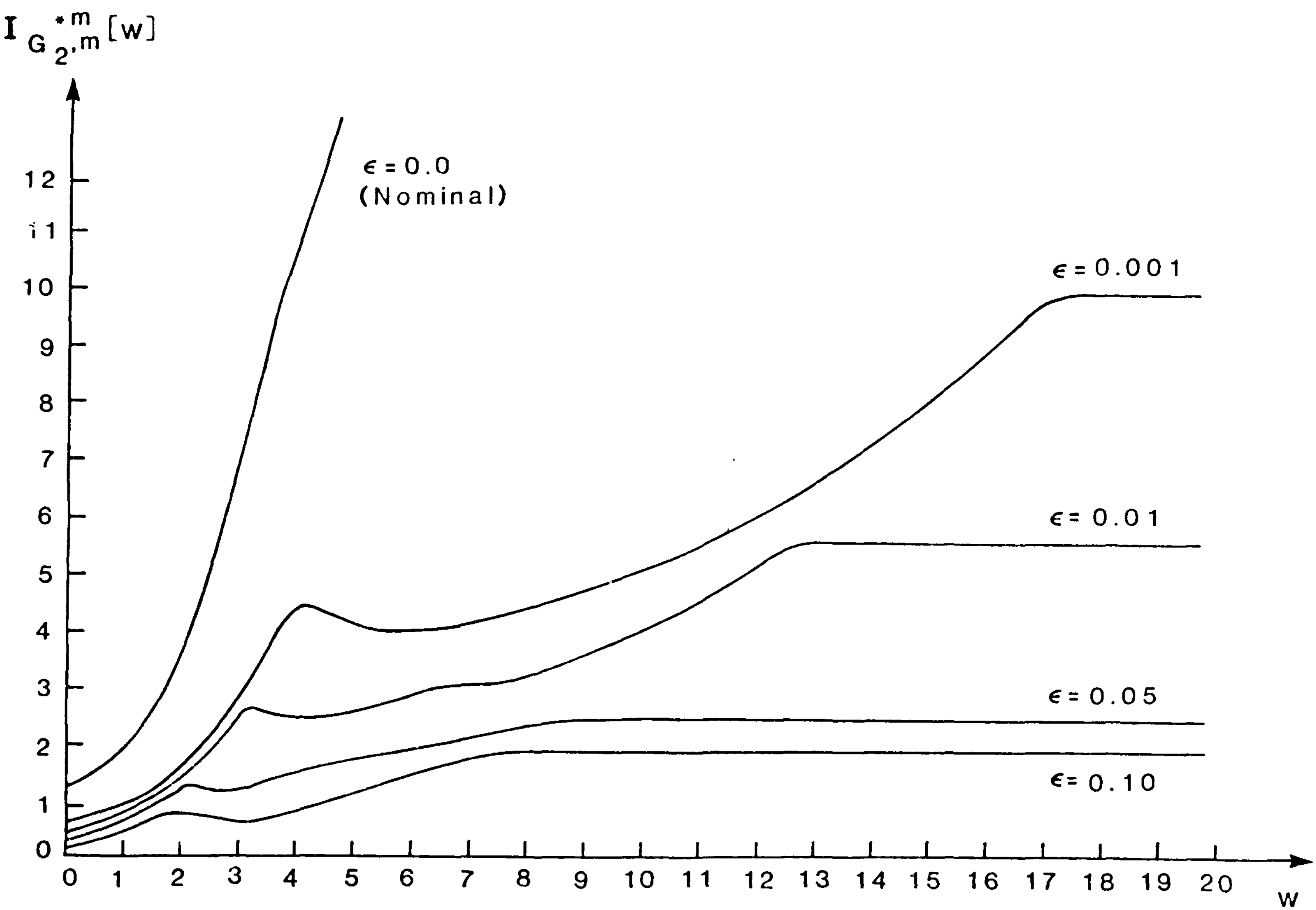


Figure 8. Influence functions $I_{G_2^*m,m}[w]$ in (96) of the nominal process 4 and m=.

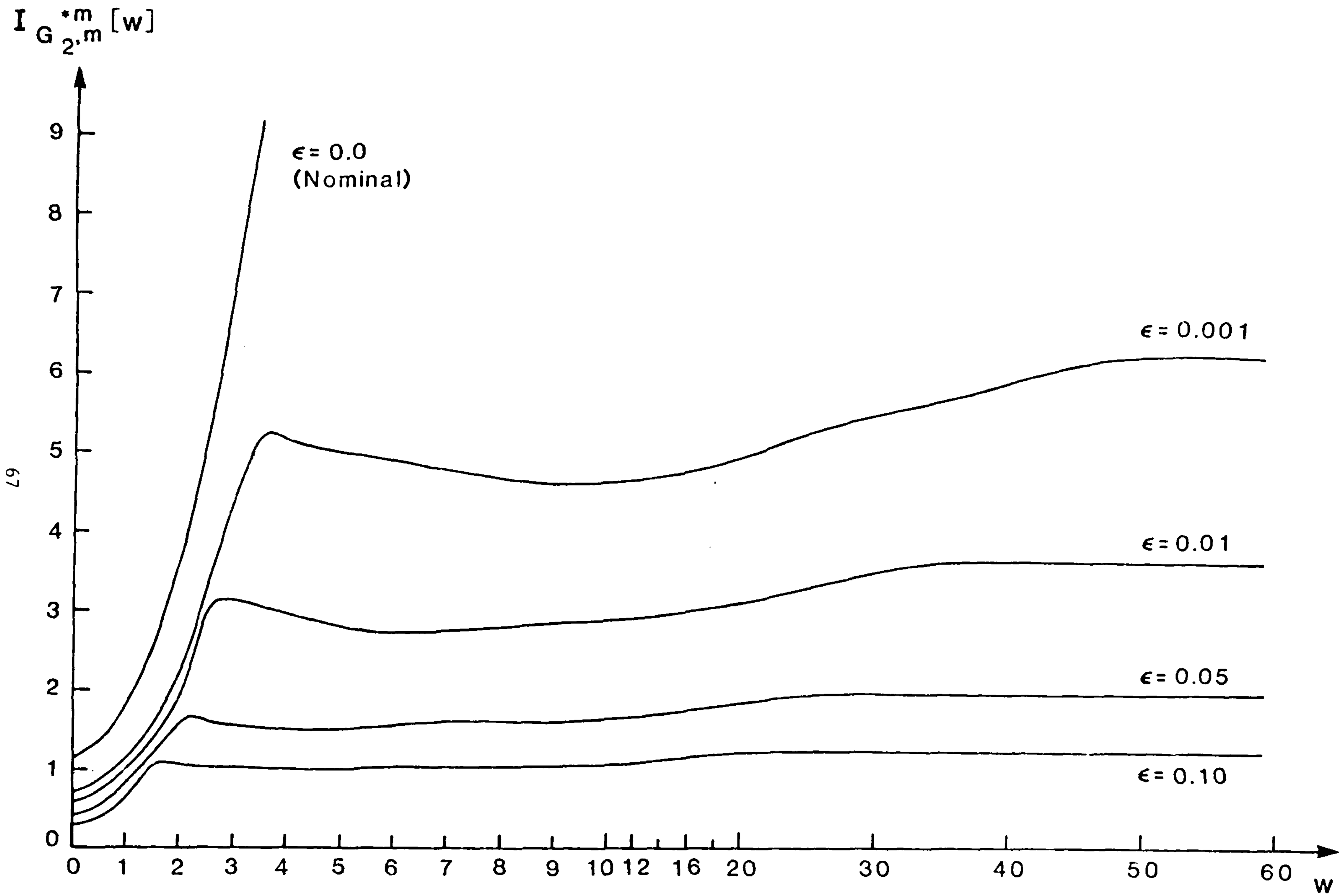


Figure [illegible]. Influence functions $I_{G_2^{*m},m}[w]$ in ([illegible]) for the nominal [illegible]

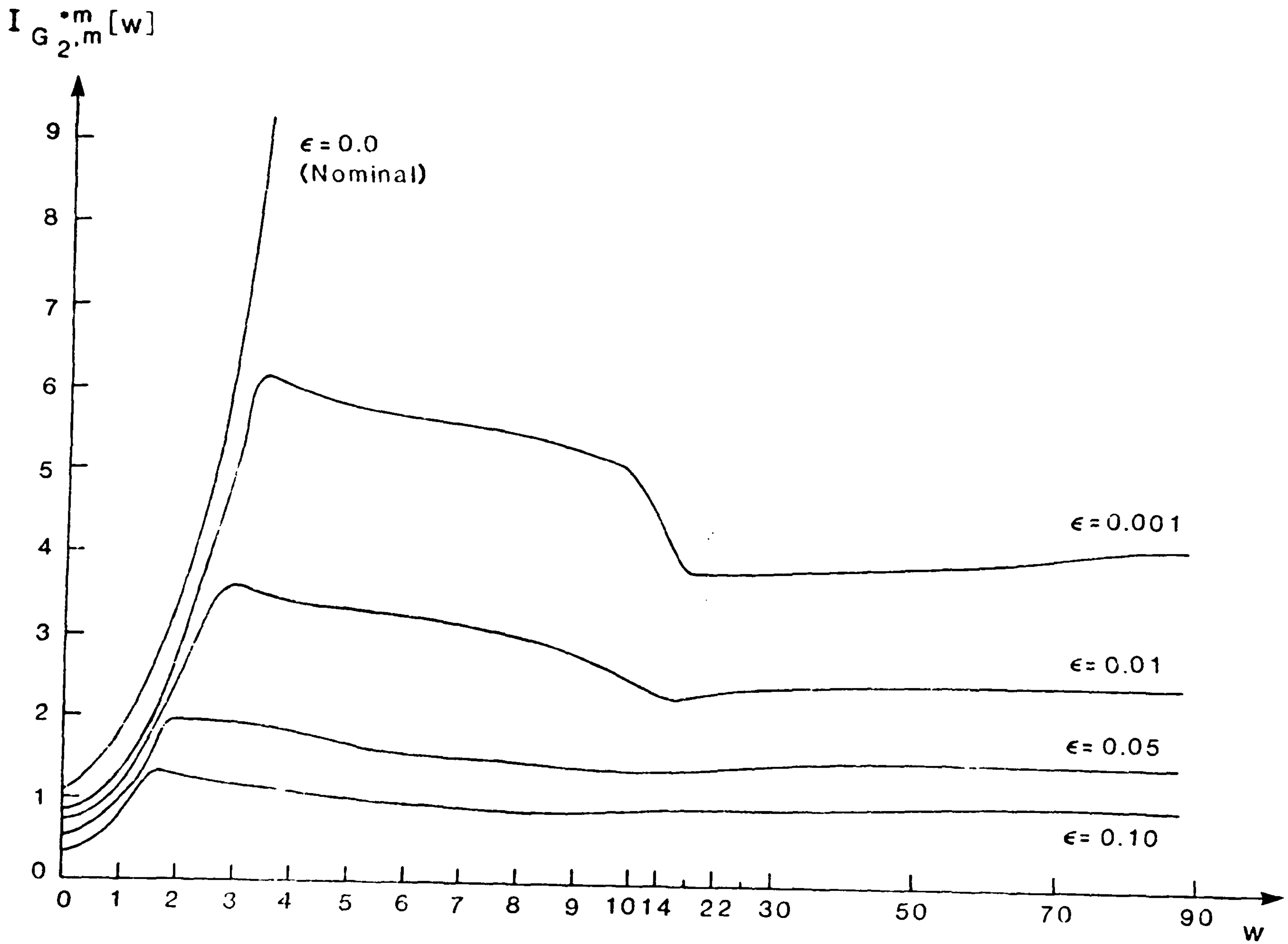


Figure 10. Influence functions $I_{G^{*m},m}[w]$ in (96) for the nominal process 4 and m=4.

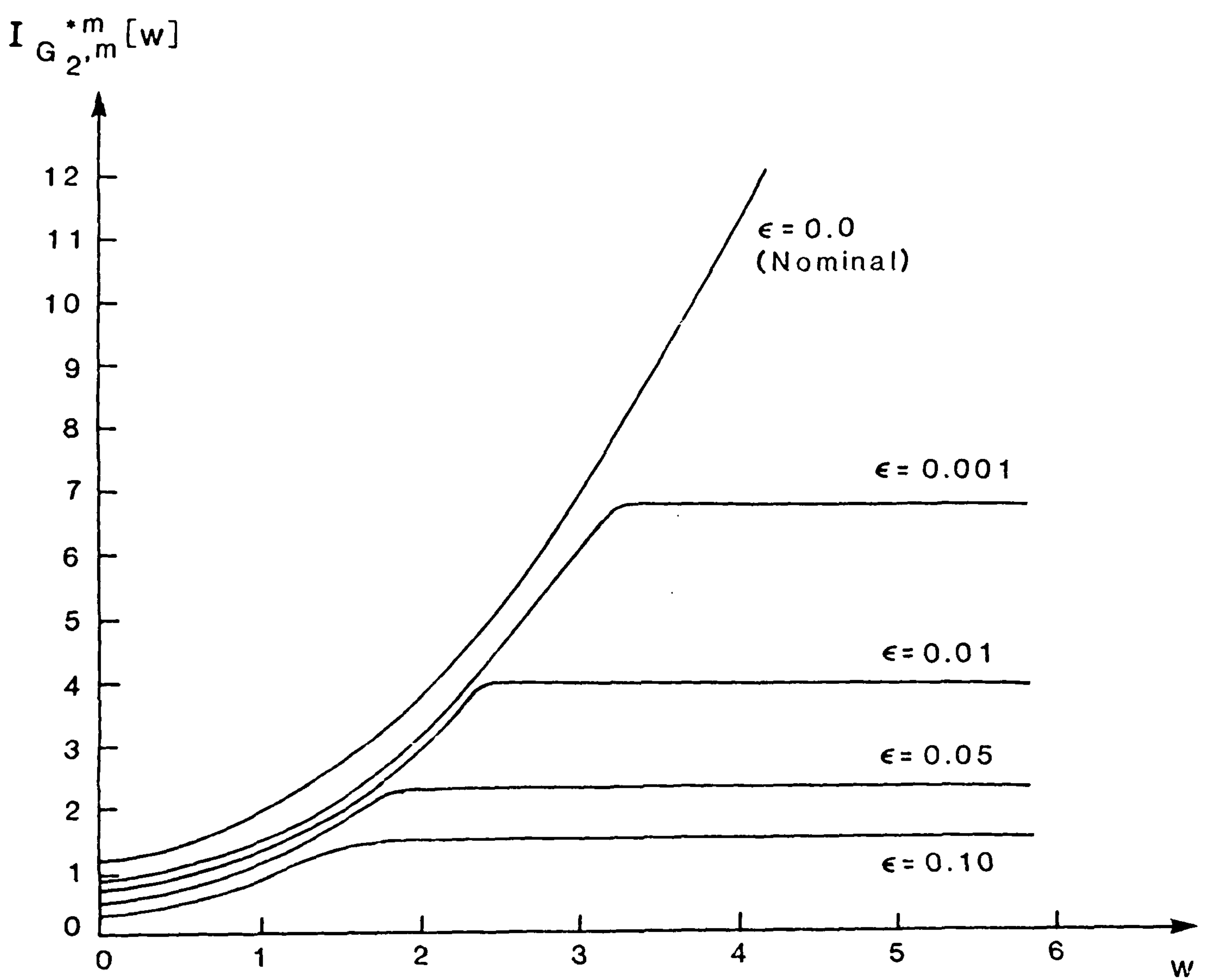


Figure 11. Influence functions $I_{G_2^{*m},m}[w]$ in [illegible] of the nominal process [illegible] m= [illegible].
For this process, $I_{G_2^{*m},m}[w]$ are identical for all $m \geq 5$.

DISTRIBUTION LIST

Copy No.	
1 - 6	Air Force Office of Scientific Research Building 410 Bolling Air Force Base Washington, D.C. 20332-6448 Attention: Major Brian W. Woodruff, NM
7 - 8	Dr. P. Papantoni-Kazakos, EE
9	Dr. Kailash Birmiwal Department of Electrical Engineering Southern illinois University Carbondale, Illinois 62901
10	Dr. R. J. Mattauch, EE
11 - 12	Ms. E. H. Pancake, Clark Hall
13	SEAS Publications Files

0685:ald:DW210R

UNIVERSITY OF VIRGINIA
School of Engineering and Applied Science

The University of Virginia's School of Engineering and Applied Science has an undergraduate enrollment of approximately 1,500 students with a graduate enrollment of approximately 560. There are 150 faculty members, a majority of whom conduct research in addition to teaching.

Research is a vital part of the educational program and interests parallel academic specialties. These range from the classical engineering disciplines of Chemical, Civil, Electrical, and Mechanical and Aerospace to newer, more specialized fields of Biomedical Engineering, Systems Engineering, Materials Science, Nuclear Engineering and Engineering Physics, Applied Mathematics and Computer Science. Within these disciplines there are well equipped laboratories for conducting highly specialized research. All departments offer the doctorate; Biomedical and Materials Science grant only graduate degrees. In addition, courses in the humanities are offered within the School.

The University of Virginia (which includes approximately 2,000 faculty and a total of full-time student enrollment of about 16,400), also offers professional degrees under the schools of Architecture, Law, Medicine, Nursing, Commerce, Business Administration, and Education. In addition, the College of Arts and Sciences houses departments of Mathematics, Physics, Chemistry and others relevant to the engineering research program. The School of Engineering and Applied Science is an integral part of this University community which provides opportunities for interdisciplinary work in pursuit of the basic goals of education, research, and public service.

Approved for public release;
distribution unlimited.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
[illegible] TRANSMITTAL TO DTIC
[illegible] report has been reviewed and is
[illegible] for public release IAW AFR 190-12.
[illegible] is unlimited.
[illegible] J. KERPER
Chief, Technical Information Division

END
DATE
FILMED
7-88
DTIC